METHODS IN MOLECULAR BIOLOGY™

Volume 298

Peptide Synthesis and Applications

Edited by

John Howl

HUMANA PRESS

Peptide Synthesis and Applications

METHODS IN MOLECULAR BIOLOGY™

John M. Walker, SERIES EDITOR

309. RNA Silencing: *Methods and Protocols,* edited by *Gordon Carmichael, 2005*

308. Therapeutic Proteins: *Methods and Protocols,* edited by *C. Mark Smales and David C. James, 2005*

307. Phosphodiesterase Methods and Protocols, edited by *Claire Lugnier, 2005*

306. Receptor Binding Techniques: *Second Edition,* edited by *Anthony P. Davenport, 2005*

305. Protein–Ligand Interactions: *Methods and Protocols,* edited by *G. Ulrich Nienhaus, 2005*

304. Human Retrovirus Protocols: *Virology and Molecular Biology,* edited by *Tuofu Zhu, 2005*

303. NanoBiotechnology Protocols, edited by *Sandra J. Rosenthal and David W. Wright, 2005*

302. Handbook of ELISPOT: Methods and Protocols, edited by *Alexander E. Kalyuzhny, 2005*

301. Ubiquitin–Proteasome Protocols, edited by *Cam Patterson and Douglas M. Cyr, 2005*

300. Protein Nanotechnology: *Protocols, Instrumentation, and Applications,* edited by *Tuan Vo-Dinh, 2005*

299. Amyloid Proteins: *Methods and Protocols,* edited by *Einar M. Sigurdsson, 2005*

298. Peptide Synthesis and Application, edited by *John Howl, 2005*

297. Forensic DNA Typing Protocols, edited by *Angel Carracedo, 2005*

296. Cell Cycle Protocols, edited by *Tim Humphrey and Gavin Brooks, 2005*

295. Immunochemical Protocols, *Third Edition,* edited by *Robert Burns, 2005*

294. Cell Migration: *Developmental Methods and Protocols,* edited by *Jun-Lin Guan, 2005*

293. Laser Capture Microdissection: *Methods and Protocols,* edited by *Graeme I. Murray and Stephanie Curran, 2005*

292. DNA Viruses: *Methods and Protocols,* edited by *Paul M. Lieberman, 2005*

291. Molecular Toxicology Protocols, edited by *Phouthone Keohavong and Stephen G. Grant, 2005*

290. Basic Cell Culture, *Third Edition,* edited by *Cheryl D. Helgason and Cindy Miller, 2005*

289. Epidermal Cells, *Methods and Applications,* edited by *Kursad Turksen, 2005*

288. Oligonucleotide Synthesis, *Methods and Applications,* edited by *Piet Herdewijn, 2005*

287. Epigenetics Protocols, edited by *Trygve O. Tollefsbol, 2004*

286. Transgenic Plants: *Methods and Protocols,* edited by *Leandro Peña, 2005*

285. Cell Cycle Control and Dysregulation Protocols: *Cyclins, Cyclin-Dependent Kinases, and Other Factors,* edited by *Antonio Giordano and Gaetano Romano, 2004*

284. Signal Transduction Protocols, *Second Edition,* edited by *Robert C. Dickson and Michael D. Mendenhall, 2004*

283. Bioconjugation Protocols, edited by *Christof M. Niemeyer, 2004*

282. Apoptosis Methods and Protocols, edited by *Hugh J. M. Brady, 2004*

281. Checkpoint Controls and Cancer, Volume 2: *Activation and Regulation Protocols,* edited by *Axel H. Schönthal, 2004*

280. Checkpoint Controls and Cancer, Volume 1: *Reviews and Model Systems,* edited by *Axel H. Schönthal, 2004*

279. Nitric Oxide Protocols, *Second Edition,* edited by *Aviv Hassid, 2004*

278. Protein NMR Techniques, *Second Edition,* edited by *A. Kristina Downing, 2004*

277. Trinucleotide Repeat Protocols, edited by *Yoshinori Kohwi, 2004*

276. Capillary Electrophoresis of Proteins and Peptides, edited by *Mark A. Strege and Avinash L. Lagu, 2004*

275. Chemoinformatics, edited by *Jürgen Bajorath, 2004*

274. Photosynthesis Research Protocols, edited by *Robert Carpentier, 2004*

273. Platelets and Megakaryocytes, Volume 2: *Perspectives and Techniques,* edited by *Jonathan M. Gibbins and Martyn P. Mahaut-Smith, 2004*

272. Platelets and Megakaryocytes, Volume 1: *Functional Assays,* edited by *Jonathan M. Gibbins and Martyn P. Mahaut-Smith, 2004*

271. B Cell Protocols, edited by *Hua Gu and Klaus Rajewsky, 2004*

270. Parasite Genomics Protocols, edited by *Sara E. Melville, 2004*

269. Vaccina Virus and Poxvirology: *Methods and Protocols,*edited by *Stuart N. Isaacs, 2004*

268. Public Health Microbiology: *Methods and Protocols,* edited by *John F. T. Spencer and Alicia L. Ragout de Spencer, 2004*

267. Recombinant Gene Expression: *Reviews and Protocols, Second Edition,* edited by *Paulina Balbas and Argelia Johnson, 2004*

266. Genomics, Proteomics, and Clinical Bacteriology: *Methods and Reviews,* edited by *Neil Woodford and Alan Johnson, 2004*

265. RNA Interference, Editing, and Modification: *Methods and Protocols,* edited by *Jonatha M. Gott, 2004*

264. Protein Arrays: *Methods and Protocols,* edited by *Eric Fung, 2004*

METHODS IN MOLECULAR BIOLOGY™

Peptide Synthesis and Applications

Edited by

John Howl

Research Institute in Healthcare Science,
School of Applied Sciences, University of Wolverhampton,
Wolverhampton, UK

HUMANA PRESS TOTOWA, NEW JERSEY

999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

www.humanapress.com

This publication is printed on acid-free paper. ∞
ANSI Z39.48-1984 (American Standards Institute)

Permanence of Paper for Printed Library Materials.

Production Editor: C. Tirpak
Cover design by Patricia F. Cleary

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: humana@humanapr.com; or visit our Website: www.humanapress.com

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

eISBN:1-59259-877-3

Library of Congress Cataloging-in-Publication Data

Peptide synthesis and applications / edited by John Howl.

p. ; cm. -- (Methods in molecular biology ; 298)

Includes bibliographical references and index.

ISBN 1-58829-317-3 (alk. paper)

1. Peptides--Synthesis--Laboratory manuals.

[DNLM: 1. Peptide Synthesis--Laboratory Manuals. QU 25 P4246 2005] I.

Howl, John. II. Series: Methods in molecular biology (Clifton, N.J.) ; v.

298.

QD431.25.S93P47 2005

547.7'56--dc22

2004020037

Preface

The broad canvas of peptide science is indebted both to its early pioneers and to the numerous international investigators who enrich this exciting discipline. The thoughts and expertise of some of these noteworthy scientists, collectively located across three continents, are represented here. The gregarious nature of the peptide science community is particularly impressive, and I am pleased that several close colleagues were able and willing to contribute to this volume. My intention, as editor of *Peptide Synthesis and Applications*, is to present the basic methodologies of contemporary peptide synthesis and to provide examples of the numerous applications that employ peptides as unique and essential materials.

As detailed in the first chapter, any reasonably competent scientist can assemble amino acids in the correct order to produce a desired peptide sequence. A course manual for basic peptide design and synthesis is also provided herein, based on a successful template used to teach peptide chemistry to undergraduate students in Stockholm. No doubt the future will see a further evolution of technologies based largely upon Merrifield's innovation of solid phase synthesis back in 1963. Thus, chapters within this volume collectively provide details of chemical ligation, the synthesis of cyclic and phosphotyrosine-containing peptides, lipoamino acid- and sugar-conjugated peptides, and more common methodologies that include peptide purification and analyses. To complete the story details of methodologies and instrumentation used for high throughput peptide synthesis are also included. Moreover, when compiling *Peptide Synthesis and Applications* my intention was to include contemporary applications of peptides that might inspire others to further expand the utility of this novel class of biomolecule. My request of many contributing authors was that they provide details of their own developments covering many different applications of peptides as novel research tools and biological probes. The design and synthesis of chimeric and cell-penetrating peptides are fields of endeavor that will no doubt provide valuable research tools and possible therapeutic leads in the foreseeable future. Details are also included of the design and application of fluorescent substrate-based peptides that can be used to determine the selectivity and activity of peptidases.

As we embrace the postgenomic era, the utility of peptides will be further exploited to both study and manipulate the many biological processes modulated by discrete molecular interactions between intracellular proteins that are a major component of the eukaryotic proteome.

Thus, *Peptide Synthesis and Applications* also includes practical details of current methodologies applicable to the identification of proteins using mass spectrometric analyses of peptide mixtures. I trust there is something here for the beginner and expert alike.

John Howl

Contents

Preface v

Contributors ix

PART I: COMMON STRATEGIES

1 Fundamentals of Modern Peptide Synthesis
Muriel Amblard, Jean-Alain Fehrentz, Jean Martinez, and Gilles Subra 3

2 Chimerism: *A Strategy to Expand the Utility and Applications of Peptides*
John Howl 25

PART II: SYNTHETIC METHODOLOGIES AND APPLICATIONS

3 Modification of Peptides and Other Drugs Using Lipoamino Acids and Sugars
Joanne T. Blanchfield and Istvan Toth 45

4 Synthesis of Linear, Branched, and Cyclic Peptide Chimera
Gábor Mezö and Ferenc Hudecz 63

5 Synthesis of Cell-Penetrating Peptides for Cargo Delivery
Margus Pooga and Ülo Langel 77

6 Incorporation of the Phosphotyrosyl Mimetic 4(Phosphonodifluoromethyl)phenylalanine (F_2Pmp) Into Signal Transduction-Directed Peptides
Zhu-Jun Yao, Kyeong Lee, and Terrence R. Burke, Jr. 91

7 Expressed Protein Ligation for Protein Semisynthesis and Engineering
Zuzana Machova and Annette G. Beck-Sickinger 105

8 Cellular Delivery of Peptide Nucleic Acid by Cell-Penetrating Peptides
Kalle Kilk and Ülo Langel 131

9 Quenched Fluorescent Substrate-Based Peptidase Assays
Rebecca A. Lew, Nathalie Tochon-Danguy, Catherine A. Hamilton, Karen M. Stewart, Marie-Isabel Aguilar, and A. Ian Smith 143

10 A Convenient Method for the Synthesis of Cyclic Peptide Libraries
Gregory T. Bourne, Jonathon L. Nielson, Justin F. Coughlan, Paul Darwen, Marc R. Campitelli, Douglas A. Horton, Andreas Rhümann, Stephen G. Love, Tran T. Tran, and Mark L. Smythe ***151***

11 High-Throughput Peptide Synthesis
Michal Lebl and John Hachmann ***167***

12 Backbone Amide Linker Strategies for the Solid-Phase Synthesis of C-Terminal Modified Peptides
Jordi Alsina, Steven A. Kates, George Barany, and Fernando Albericio ***195***

13 Synthesis of Peptide Bioconjugates
Ferenc Hudecz ***209***

Part III: Practical Guides

14 Protein Identification by Mass Spectrometric Analyses of Peptides
Ashley Martin ***227***

15 Manual Solid-Phase Synthesis of Glutathione Analogs: *A Laboratory-Based Short Course*
Ursel Soomets, Mihkel Zilmer, and Ülo Langel ***241***

Index 259

Contributors

MARIE-ISABEL AGUILAR • *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia*
FERNANDO ALBERICIO • *Barcelona Biomedical Research Institute, Barcelona Science Park, University of Barcelona, Barcelona, Spain*
JORDI ALSINA • *Eli Lily and Company, Indianapolis, IN*
MURIEL AMBLARD • *Laboratoire des Amino Acides, Peptides et Protéines-UMR-CNRS 5810, Faculté de Pharmacie, Montpellier, France*
GEORGE BARANY • *Department of Chemistry, University of Minnesota, Minneapolis, MN*
ANNETTE G. BECK-SICKINGER • *Institute of Biochemistry, Faculty of Biosciences, Pharmacy, and Psychology, University of Leipzig, Leipzig, Germany*
JOANNE T. BLANCHFIELD • *School of Molecular and Microbial Sciences, University of Queensland, St. Lucia, Queensland, Australia*
GREGORY T. BOURNE • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
TERRENCE R. BURKE, JR. • *Laboratory of Medicinal Chemistry, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD*
MARC R. CAMPITELLI • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
JUSTIN F. COUGHLAN • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
PAUL DARWEN • *Protagonist Pty. Ltd., Queensland Bioscience Precinct, University of Queensland, St. Lucia, Queensland, Australia*
JEAN-ALAIN FEHRENTZ • *Laboratoire des Amino Acides, Peptides et Protéines-UMR-CNRS 5810, Faculté de Pharmacie, Montpellier, France*
JOHN HACHMANN • *Illumina Inc., San Diego, CA*
CATHERINE A. HAMILTON • *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia*
DOUGLAS A. HORTON • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
JOHN HOWL • *Research Institute in Healthcare Science, School of Applied Sciences, University of Wolverhampton, Wolverhampton, UK*
FERENC HUDECZ • *Research Group of Peptide Chemistry, Department of Organic Chemistry, Hungarian Academy of Sciences, Eötvös Loránd University, Budapest, Hungary*
STEVEN A. KATES • *CereMedix Inc., Maynard, MA*
KALLE KILK • *Department of. Neurochemistry and Neurotoxicology, University of Stockholm, Stockholm, Sweden.*

Ülo Langel • *Department of Neurochemistry and Neurotoxicology, Arrhenius Laboratories, Stockholm University, Stockholm, Sweden*
Michal Lebl • *Illumina Inc., San Diego, CA*
Kyeong Lee • *Laboratory of Medicinal Chemistry, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD*
Rebecca A. Lew • *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia*
Stephen G. Love • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
Zuzana Machova • *Institute of Biochemistry, Faculty of Biosciences, Pharmacy, and Psychology, University of Leipzig, Leipzig, Germany*
Ashley Martin • *Cancer Research UK Institute for Cancer Studies, University of Birmingham, Birmingham, UK*
Jean Martinez • *Laboratoire des Amino Acides, Peptides et Protéines-UMR-CNRS 5810, Faculté de Pharmacie, Montpellier, France*
Gábor Mezö • *Research Group of Peptide Chemistry, Hungarian Academy of Sciences, Eötvös Loránd University, Budapest, Hungary*
Jonathon L. Nielson • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
Margus Pooga • *Estonian Biocentre, Tartu, Estonia*
Andreas Rhümann • *Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
A. Ian Smith • *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia*
Mark L. Smythe • *Protagonist Pty. Ltd., Queensland Bioscience Precinct, Institute for Molecular Bioscience, University of Queensland, St. Lucia, Queensland, Australia*
Ursel Soomets • *Department of Biochemistry, Tartu University, Tartu, Estonia*
Karen M. Stewart • *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria, Australia*
Gilles Subra • *Laboratoire des Amino Acides, Peptides et Protéines-UMR-CNRS 5810, Faculté de Pharmacie, Montpellier, France*
Nathalie Tochon-Danguy • *Department of Pharmaceutical Biology and Pharmacology, Monash University, Parkville, Victoria, Australia*
Istvan Toth • *School of Pharmacy, University of Queensland, St. Lucia, Queensland, Australia*
Tran T. Tran • *Protagonist Pty. Ltd, Queensland Bioscience Precinct, University of Queensland, St. Lucia, Queensland, Australia*
Zhu-Jun Yao • *Laboratory of Medicinal Chemistry, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD*
Mihkel Zilmer • *Department of Biochemistry, Tartu University, Tartu, Estonia*

I

Common Strategies

1

Fundamentals of Modern Peptide Synthesis

Muriel Amblard, Jean-Alain Fehrentz, Jean Martinez, and Gilles Subra

Summary

The purpose of this chapter is to delineate strategic considerations and provide practical procedures to enable non-experts to synthesize peptides with a reasonable chance of success. This chapter focuses on Fmoc chemistry, which is now the most commonly employed strategy for solid phase peptide synthesis (SPPS). Protocols for the synthesis of fully deprotected peptides are presented, together with a review of linkers and supports currently employed for SPPS. The principles and the different steps of SPPS (anchoring, deprotection, coupling reaction, and cleavage) are all discussed, along with their possible side reactions.

Key Words: Solid phase peptide synthesis; side reaction coupling; anchoring; deprotection; cleavage; linker.

1. Introduction

Nowadays, "peptide synthesis" includes a large range of techniques and procedures that enable the preparation of materials ranging from small peptides to large proteins. The pioneering work of Bruce Merrifield ***(1)***, which introduced solid phase peptide synthesis (SPPS), dramatically changed the strategy of peptide synthesis and simplified the tedious and demanding steps of purification associated with solution phase synthesis. Moreover, Merrifield's SPPS also permitted the development of automation and the extensive range of robotic instrumentation now available. After defining a synthesis strategy and programming the amino acid sequence of peptides, machines can automatically perform all the synthesis steps required to prepare multiple peptide samples. SPPS has now become the method of choice to produce peptides, though solution phase synthesis can still be useful for large-scale production of a given peptide.

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

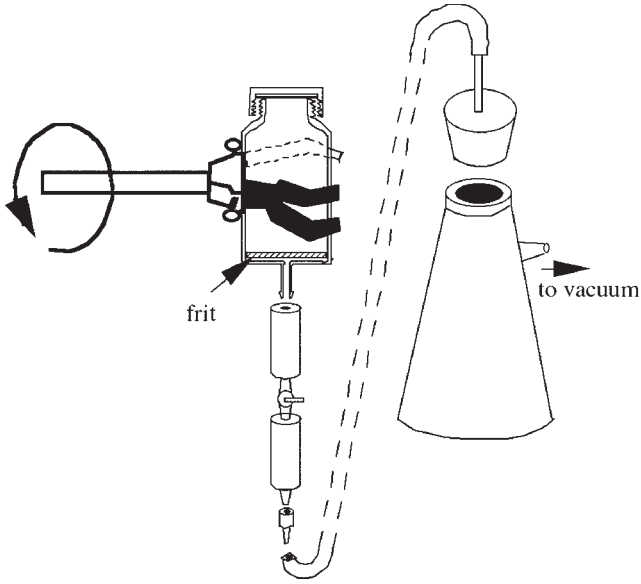


Fig. 1. Basic equipment for SPPS.

2. Materials

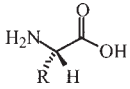
1. Reaction vessel (**Fig. 1**).
2. Polytetrafluoroethylene (PTFE) stick (15 cm length, 0.6–0.8 cm diameter).
3. Rotor.
4. Filtration flask.
5. Porous frit.
6. Lyophilizer.
7. HPLC equipped with reverse phase C_{18} column.
8. pH-Indicating paper.
9. Solvents (*N,N*-dimethylformamide [DMF], methanol [MeOH], dichloromethane [DCM]) in wash bottles.
10. Diisopropylethylamide (DIPEA).
11. Piperidine solution in DMF (20:80).
12. Kaiser test solutions (ninhydrin, pyridine, phenol) (*see* **Note 1**).
13. Fmoc-amino-acids with protected side-chains (*see* **Table 1**).
14. Trifluoroacetic acid (TFA).
15. Triisopropylsilane (TIS).
16. *tert*-butyl methyl ether (MTBE).

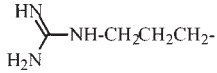
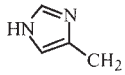
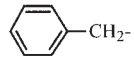
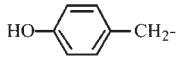
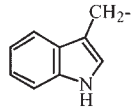
3. Methods

3.1. Principles of SPPS

As peptide synthesis involves numerous repetitive steps, the use of a solid support has obvious advantages. With such a system a large excess of reagents at

Table 1
Proteinogenic Amino Acids



Amino acid	Three-letter code	One-letter code	Side-chain
Glycine	Gly	G	H-
Alanine	Ala	A	CH_3-
Valine	Val	V	$(CH_3)_2CH$-
Leucine	Leu	L	$(CH_3)_2CHCH_2$-
Isoleucine	Ile	I	$CH_3CH_2(CH_3)CH$-
Aspartic acid	Asp	D	$HOOC\text{-}CH_2$-
Asparagine	Asn	N	$H_2NOC\text{-}CH_2$-
Glutamic acid	Glu	E	$HOOC\text{-}CH_2CH_2$-
Glutamine	Gln	Q	$H_2NOC\text{-}CH_2CH_2$-
Lysine	Lys	K	$H_2N\text{-}CH_2CH_2CH_2CH_2$-
Arginine	Arg	R	
Histidine	His	H	
Serine	Ser	S	$HO\text{-}CH_2$-
Threonine	Thr	T	$CH_3\text{-}CH(OH)$-
Phenylalanine	Phe	F	
Tyrosine	Tyr	Y	
Tryptophan	Trp	W	
Cysteine	Cys	C	$HS\text{-}CH_2$-
Methionine	Met	M	$CH_3\text{-}S\text{-}CH_2CH_2$-
Proline	Pro	P	

Note: All the 20 DNA-encoded or proteinogenic α-amino acids are of L stereochemistry.

high concentration can drive coupling reactions to completion. Excess reagents and side products can be separated from the growing and insoluble peptide simply by filtration and washings, and all the synthesis steps can be performed in the same vessel without any transfer of material.

The principles of SPPS are illustrated in **Fig. 2**. The *N*-protected C-terminal amino acid residue is anchored via its carboxyl group to a hydroxyl (or chloro)

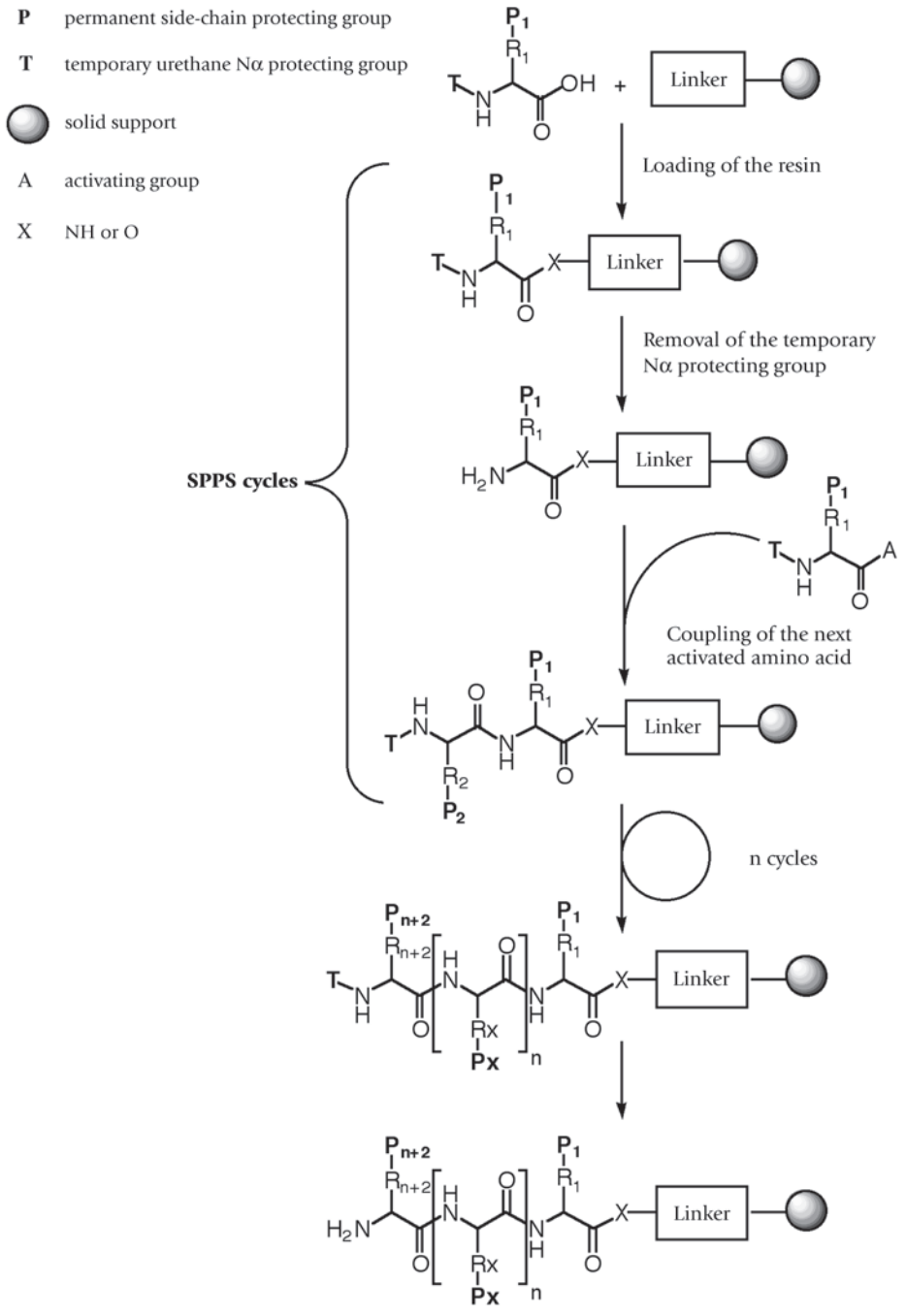


Fig. 2. Principles of SPPS.

or amino resin to yield respectively an ester or amide linked peptide that will ultimately produce a C-terminal acid or a C-terminal amide peptide. After loading the first amino acid, the desired peptide sequence is assembled in a linear fashion from the C-terminus to the N-terminus (the C → N strategy) by repetitive cycles of N^{α} deprotection and amino acid coupling reactions.

Side-chain functional groups of amino acids must be masked with permanent protecting groups (P_n) that are stable in the reaction conditions used during peptide elongation. The α-amino group is protected by a temporary protecting group (T) that is usually a urethane derivative. The temporary protecting group (T) can be easily removed under mild conditions that preserve peptide integrity and reduce the rate of epimerization, which can occur via 5(4*H*)-oxazolone formation of the activated amino acid during the coupling step ***(2,3)*** as indicated in **Fig. 3**. The protective role of urethanes against epimerization also explains the predominance of the C → N strategy.

After coupling, the excess of reactants is removed by filtration and washings. The temporary N-terminal protecting group is removed allowing the addition of the next *N*-urethane protected amino acid residue by activation of its α-carboxylic acid. This process (deprotection/coupling) is repeated until the desired sequence is obtained. In a final step, the peptide is released from the resin and the side-chain protecting groups (P_n) concomitantly removed.

3.2. Fmoc/tBu SPPS

In SPPS, two main strategies are used: the Boc/Bzl and the Fmoc/tBu approaches for T/P_n protecting groups. The former strategy is based on the graduated acid lability of the side-chain protecting groups. In this approach, the Boc group is removed by neat TFA or TFA in dichloromethane, and side-chain protecting groups and peptide-resin linkages are removed at the end of the synthesis by treatment with a strong acid such as anhydrous hydrofluoric acid (HF). While this method allows efficient syntheses of large peptides and small proteins, the use of highly toxic HF and the need for special polytetrafluoroethylene-lined apparatus limit the applicability of this approach to specialists only. Moreover, the use of strongly acidic conditions can produce deleterious changes in the structural integrity of peptides containing fragile sequences.

The Fmoc/*t*Bu method ***(4)*** is based on an orthogonal protecting group strategy. This approach uses the base-labile *N*-Fmoc group for protection of the α-amino function, acid-labile side-chain protecting groups and acid-labile linkers that constitute the C-terminal amino acid protecting group. This latter strategy has the advantage that temporary and permanent orthogonal protections are removed by different mechanisms, allowing the use of milder acidic conditions for final deprotection and cleavage of the peptide from the resin. For all these reasons, Fmoc-based SPPS is now the method of choice for the routine synthesis of peptides.

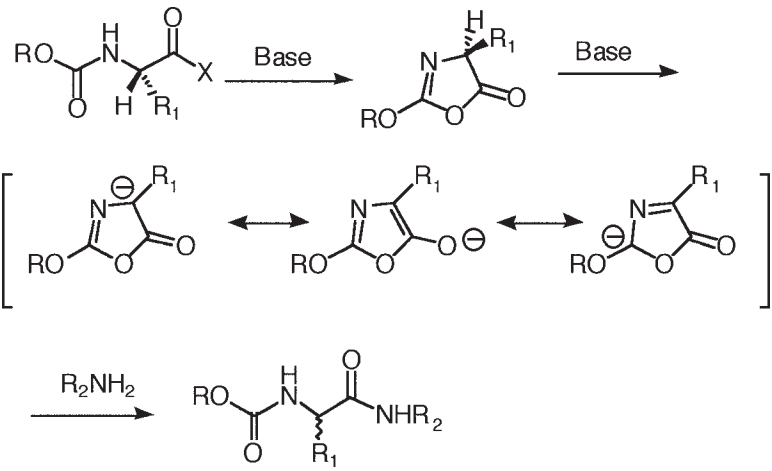


Fig. 3. Epimerization by oxazolone formation.

3.3. Solid Supports

The matrix polymer and the linker can characterize solid supports ***(5)***. Often the term "resin" is improperly used in place of the linker system, ignoring the fact that the matrix polymer is as important in supported chemistry as the solution phase ***(6)***. As hundreds of different resins are commercially available, some of them carrying the same linkers, special care should be taken to properly choose the most suitable linker for the synthesis.

3.3.1. Matrix Polymers

Cross-linked polystyrene (PS)-based resins are most commonly used for routine SPPS. Beads of 200 to 400 mesh size distribution (corresponding to a diameter of about 50 μm) and a loading of 0.5 to 0.8 mmol/g present good characteristics for polymer swelling in solvents such as DMF and DCM, diffusion of reactants into the polymer matrix, and accessibility of linker sites buried into the bead. For larger peptides (more than 25 amino acids) or more difficult sequences, a lower loading is required (0.1–0.2 mmol/g).

Cross-linked polyamide (PA)-based resins and composite PS-polyethylene glycol (PEG)-based resins are much more hydrophilic supports exhibiting physical properties different from PS resins at microscopic and macroscopic levels ***(7)***. These supports, often with a lower loading capacity, may represent an alternative to standard cross-linked PS resins for the synthesis of difficult sequences and large peptides.

Table 2
Type of SPPS Reaction Vessels

Vessel length (cm)	Vessel diameter (cm)	Max. resin weight (g)	Working volume (mL)
5	2	0.5	10
11	2.6	2	40
15	3.4	4	90

3.4. Resin Handling

3.4.1. SPPS Reaction Vessels

SPPS can be performed in classical glass reaction vessels that can be made by glassblowers or purchased from manufacturers (**Fig. 1**). Alternatively, syringes equipped with PTFE or glass frits may also be used.

Reaction vessel size should be in relation to the amount of resin used, according to **Table 2**.

3.4.2. Solvents

As 99% of coupling sites are not at the surface but inside the resin beads, swelling of beads carrying the growing peptide chain is essential for the optimal permeation of activated *N*-protected amino acids within the polymer matrix, thus improving coupling yields. Before starting the solid phase synthesis, the resin has to be swollen in an adequate solvent such as DCM or DMF for 20 to 30 min (**Protocol 1**). For cross-linked polystyrene beads used in SPPS, DCM presents optimal swelling properties. For coupling steps, polar aprotic solvents such as DMF or *N*-methylpyrrolidone (NMP) are preferred to improve solubility of reactants. Alcohols and water are not adequate solvents for PS resins (*see* **Note 2**). Nevertheless, methanol or isopropanol can be used during the washing steps (**Protocol 2**) to shrink PS resin beads. This shrinking will efficiently remove reactants in excess. After such treatment, PS beads should be swollen in DCM or DMF.

3.4.3. Stirring and Mixing

It is not necessary to agitate the reaction vessel vigorously, as diffusion phenomena dictate the kinetic reaction in SPPS. Moreover, most types of resin beads used for peptide synthesis are fragile, so magnetic stirring is not recommended. An old rotary evaporator rotor can be used for stirring during coupling and deprotection steps or alternately any apparatus enabling smooth agitation by rocking

or vortexing is appropriate. As beads usually stick to glass, the important condition is that all surfaces of the reactor must be in contact with the reaction mixture during stirring.

3.4.4. Washing

Washing steps are essential to remove soluble side products and the excess of reactants used during coupling and deprotection steps. Filling the reactor with solvent contained in a wash bottle and emptying it under vacuum is an appropriate and simple method. If necessary, stirring and mixing of the resin in the washing solvent can be performed with a PTFE stick.

Protocol 1. Resin Swelling

1. Place the dry resin in the appropriate reaction vessel (*see* **Subheading 3.4.1.**).
2. Fill the reactor with DCM until all resin beads are immersed. Resin suspension can be gently mixed with a PTFE stick.
3. Leave for 20 to 30 min.
4. Remove DCM by filtration under vacuum.

Protocol 2. Standard Washing Procedures

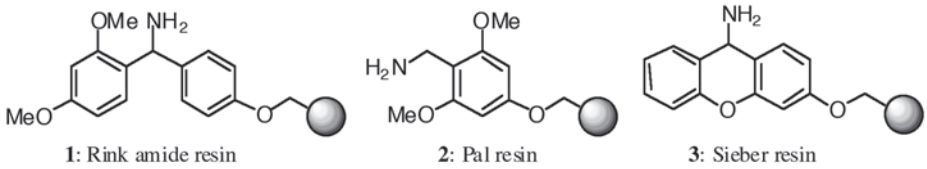
1. Fill the reaction vessel with DMF.
2. Leave for 10 s and remove the solvent by filtration.
3. Carefully wash the screw cap and the edge of the reactor with DMF.
4. Repeat **steps 1** and **2** twice with DMF.
5. Repeat **steps 1** and **2** with MeOH.
6. Repeat **steps 1** and **2** with DCM.
7. Repeat **steps 1** and **2** with DMF.

3.5. Linkers and Resins for Fmoc-Based SPPS

The first step of SPPS is the anchoring (or loading) of the *N*-protected C-terminal amino acid residue to the solid support via an ester or an amide bond depending on the C-terminal functional group of target peptide (respectively acid or amide). Most of the linkers are commercially available anchored on the different matrices (PS, PA, PEG-PS). The bead (ball) symbol used in the following paragraph is generic and does not refer to a particular matrix.

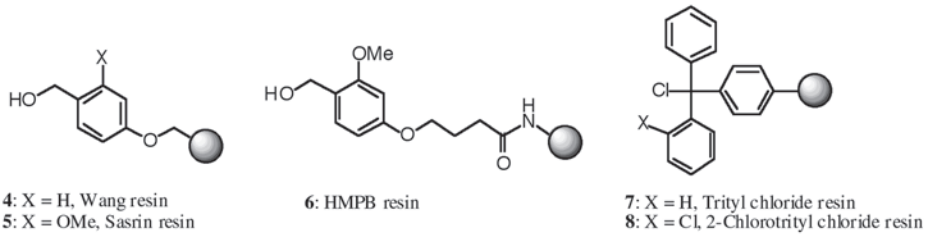
3.5.1. Peptide Amides

For the synthesis of C-terminal peptide amides, the commonly used resins are 1 to 3 (**Fig. 4**) ***(8–10)***. These resins are compatible with Fmoc chemistry and final TFA cleavage. For attachment of the first urethane *N*-protected residue, standard peptide coupling procedures (**Protocol 5**) can be used. These resins are usually supplied Fmoc-protected and should be deprotected before incorpora-



1: Rink amide resin **2**: Pal resin **3**: Sieber resin

Fig. 4. Resins for peptide amide synthesis.



4: X = H, Wang resin
5: X = OMe, Sasrin resin

6: HMPB resin

7: X = H, Trityl chloride resin
8: X = Cl, 2-Chlorotrityl chloride resin

Fig. 5. Resins for peptide acid synthesis.

tion of the first residue. With bulky C-terminal amino acids, a double coupling step can be necessary.

3.5.2. Peptide Acids

The anchoring of an amino acid to the solid support by esterification is often more difficult, and even hazardous, for some residues and can lead to epimerization, dipeptide formation, and low substitution. Thus, we recommend the purchase of resins preloaded with the first C-terminal *N*-protected amino acid; these are commercially available from various manufacturers.

Commonly used resins in Fmoc/tBu strategy for the synthesis of C-terminal peptide acid are reported in **Fig. 5** ***(11–14)***. Anchoring reactions must be performed in an anhydrous medium and amino acids containing water should be dried before use.

3.5.3. Hydroxymethyl-Based Resins

For hydroxymethyl-based resins 4 to 6, formation of the ester linkage is easier with unhindered resins such as Wang resin 4 compared with resins possessing withdrawing methoxy groups 5 and 6. The most commonly used esterification process is the symmetrical anhydride method (**Protocol 3**). Determination of the loading can be performed by Fmoc release measurement (*see* **Note 3**). In the case of difficult anchoring, the esterification step can be repeated with fresh reactants. Arginine derivatives can need three esterification steps to achieve

correct loading. After anchoring, unreacted resin-bound hydroxyl groups should be capped by benzoic anhydride or acetic anhydride (*see* **Protocol 3, step 8**).

PROTOCOL 3. ATTACHMENT TO HYDROXYMETHYL-BASED RESIN

1. Place the resin in the appropriate SPPS reactor.
2. Swell the resin as described in **Protocol 1**.
3. The desired Fmoc amino acid (10 eq relative to resin substitution) is placed in a dry, round-bottom flask with a magnetic stirrer and dissolved in dry DCM at 0°C (3 mL/mmol). Some drops of dry DMF may be useful to achieve complete dissolution.
4. Add DIC (5 eq) and stir the mixture for 10 min at 0°C. If a precipitate is observed, add DMF until dissolution and stir for 10 min longer.
5. Add the solution to the hydroxylmethyl resin.
6. Dissolve dimethylaminopyridine (DMAP) (0.1–1 eq) in DMF and add the solution to the reaction mixture.
7. After 1 h stirring, wash the resin with DMF (three times) and finally with DCM.
8. Dry the resin *in vacuo* for 18 h before performing Fmoc release measurement on a sample (*see* **Note 3**). When the loading is less than 70% repeat the esterification step.
9. When the desired substitution is achieved, cap the remaining hydroxyl groups by adding benzoic or acetic anhydride (5 eq) and pyridine (1 eq) in DMF to the resin (previously swelled) and stir for 30 min.
10. Wash the resin (**Protocol 2**) and start classical elongation with *N*-protected amino acid after Fmoc deprotection (**Protocol 7**).

3.5.4. Trityl-Based Resins

Trityl-based resins are highly acid-labile. The steric hindrance of the linker prevents diketopiperazine formation and the resins are recommended for Pro and Gly C-terminal peptides. Extremely mild acidolysis conditions enable the cleavage of protected peptide segments from the resin. These resins are commercially available as their chloride or alcohol precursors. The trityl chloride resin is extremely moisture-sensitive, so reagents and glassware should be carefully dried before use to avoid hydrolysis into the alcohol form. It is necessary to activate the trityl alcohol precursor and it is highly recommended to reactivate the chloride just before use (*see* **Note 4**). After activation, attachment of the first residue occurs by reaction with the Fmoc amino acid derivative in the presence of a base. This reaction does not involve an activated species, so it is free from epimerization. Special precautions should be taken for Cys and His residues that are particularly sensitive to epimerization during activation (**Table 2**).

PROTOCOL 4. ATTACHMENT TO TRITYL-BASED RESIN

1. Place 1 g of trityl-based resin (1.0–2.0 mmol chloride/g resin) in an SPPS reaction vessel.

2. Swell the resin as described in **Protocol 1**.
3. Add a solution of 3 eq of Fmoc amino acid and 7.5 eq of DIPEA in dry DCM (10 mL/g resin). When a lower substitution resin is desired, reduce the amount of amino acid.
4. Stir the mixture for 30 to 60 min at room temperature.
5. Wash the resin with DMF (two to three times).
6. Add 10 mL of a mixture of DCM/MeOH/DIPEA (80:15:5) to cap any remaining reactive chloride group.
7. Mix for 15 min and filter.
8. Wash the resin three times with DMF and DCM. After drying *in vacuo*, the substitution can be measured from Fmoc release (*see* **Note 3**).

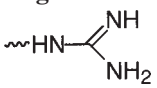
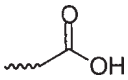
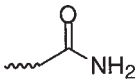
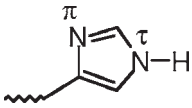
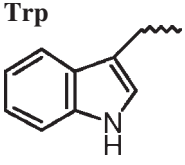
3.6 Side-Chain Protecting Groups

We will limit the description of side-chain protecting groups ***(5)*** to those that have been found most effective for the preparation of a large number of classical peptide sequences in Fmoc SPPS and are commercially available from most of the protected amino acid providers. For routine synthesis, TFA-labile protecting groups are usually used. However, for selective modifications of a particular residue on the solid support (e.g., side-chain cyclized peptides, biotinylated peptides), special orthogonal protecting groups are needed ***(15)***. Some of the most commonly used side-chain protecting groups are reported in **Table 3**.

3.7. Coupling Reaction

The most simple and rapid procedure for the stepwise introduction of *N*-protected amino acids in SPPS is the *in situ* carboxylic function activation (**Protocol 5**). A large excess of the activated amino acid is used (typically 2–10 times excess compared to the resin functionality, which is provided by the manufacturer or empirically determined; *see* **Note 3**). This excess allows a high concentration of reactants (typically 60–200 m*M*) to ensure effective diffusion. The time required for a complete acylation reaction depends on the nature of the activated species, the peptide sequence that is already linked to the resin, and the concentration of reagents. This last parameter must be as high as possible and is in connection with the volume of the reaction vessel and the resin substitution (**Table 2**). The preferred coupling reagents for *in situ* activation are benzotriazol-1-yl-oxytripyrrolidino phosphonium hexafluorophosphate (PyBOP) ***(16)*** for phosphonium-based activation and *O*-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) ***(17)*** or *N*-[1H-benzotriazol-1-yl) (dimeth-ylamino)methylene]-*N*-methyl-methanaminium hexafluorophosphate *N*-oxide (HBTU) ***(18)*** for aminium/uronium-based activation (**Fig. 6**). These coupling reagents convert *N*-protected amino acids into their corresponding OBt esters. A tertiary amine (generally diisopropylethylamine) is required to produce

Table 3
Side-Chain Protecting Groups in Fmoc-Based SPPS

Amino acid and side-chain functionality	Protecting groups	Removal condition	Remarks and side reactions
Arg 	Pmc Pbf	95% TFA 95% TFA	Presence of thiols may accelerate the cleavage.
Asp/Glu 	OtBu OAll[a]	95% TFA $Pd(Ph_3P)_4/PhSiH_3$	Aspartimide formation (*see* **Subheading 12.2.**).
Asn/Gln 	Trt	95% TFA	Protections avoid dehydration of the carboxamide side-chain during activation and help to solubilize Fmoc-Asn-OH and Fmoc-Gln-OH. pyrGlu formation for N-terminal glutamine peptides (*see* **Note 5**).
Cys 	Trt	95% TFA	High level of epimerization can occur during activation (*see* **Note 6**). Participate in the folding of peptides and proteins by disulfide bridge formation (*see* **Subheading 3.11.**).
His 	Trt (NHτ) Mtt[a]	TFA 1% TFA	Even with protection of the imidazole ring, problems of epimerization can occur during activation.
Lys 	Boc Mtt[a] Aloc[a]	TFA 1% TFA $Pd(Ph_3P)_4/PhSiH_3$	
Ser/Thr/Tyr 	tBu Trt[a]	TFA 1% TFA	
Trp 	Boc	TFA	Trp can be used unprotected. However, if Arg(Pmc) or Arg(Pbf) is present in the sequence, side-reaction can occur.

[a]Protection used for on-resin derivatization.

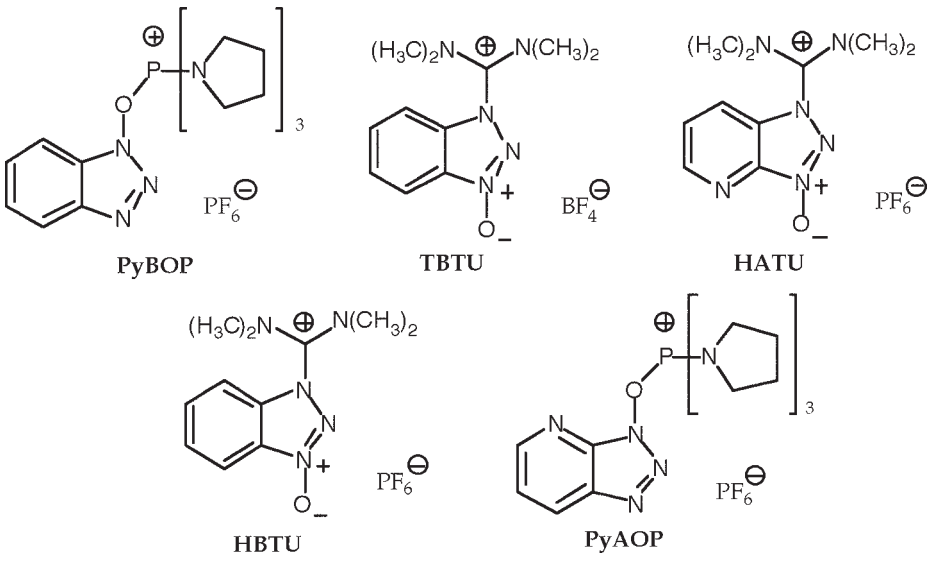


Fig. 6. Phosphonium and uronium/aminium coupling reagents.

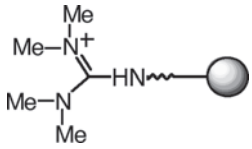


Fig. 7. N-terminal tetramethylguanidinated peptides.

the carboxylate of *N*-protected amino acids which reacts with coupling reagents. More recently, *N*-[1H-benzotriazol-1-yl(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorophosphate (HATU) ***(19)*** and 7-azabenzotriazol-1-yl-oxytris(pyrrolidino)phosphonium hexafluorophosphate (PyAOP) ***(20)***, that generate OAt esters, have been reported to be more efficient and to reduce epimerization.

Uronium/aminium-based reagents (HBTU, HATU, TBTU) are thought to mechanistically function in a similar way to their phosphonium analogs but unlike them, they can irreversibly block the free N-terminal amino function of the peptide-resin by forming tetramethylguanidinium derivatives (**Fig. 7**) ***(21)***. To avoid this side-reaction, it is recommended to generate the carboxylate of the amino acid before adding these coupling reagents at a slightly reduced equivalent compared to the amino acid.

Protocol 5. Standard Coupling Procedure With HBTU

1. After Fmoc deprotection and washings (or properly conditioning and swelling when the resin is dry), wash the resin once with DMF.
2. Add the *N*-α Fmoc-protected amino acid as a powder (usually 3 eq).
3. Fill the SPPS reaction vessel (at least 2/3 of volume) with DMF and stir with a PTFE stick for 10–20 s.
4. Add DIPEA to the vessel (usually slight excess, 3.5–4 eq) and stir with a PTFE stick until complete dissolution of the *N*-protected amino acid occurs. Other tertiary amines can be used such as *N*-methyl morpholine or triethylamine.
5. Add HBTU (2.9 eq compared with 3 eq of amino acid), screw the cap, and stir for 30 min. For difficult sequences it is recommended to preactivate the *N*-protected amino acid before coupling to the free N^{α}-amino function to limit the guanylation side-reaction.
6. Remove the coupling solution by filtration.
7. Wash the resin properly according to **Protocol 2**.
8. Perform qualitative monitoring of the coupling reaction using a colorimetric test (*see* **Note 1**).

In some cases coupling is not complete and a colorimetric test will reveal the presence of free amino groups. On these occasions a double coupling with fresh reagents must be performed. When acylation is incomplete after a second coupling, a capping procedure through acetic anhydride (**Protocol 6**) can be performed to stop the elongation of these less-reactive amino groups. When the colorimetric test still reveals the presence of free amine functions after capping, aggregation can be suspected.

Protocol 6. Capping

1. Swell the resin in DCM.
2. Remove the DCM by filtration.
3. Fill the SPPS reactor (at least 2/3 of volume) with a 50/50 DCM/acetic anhydride solution (in case of trityl linker, *see* **Note 7**) and mix for 3 min.
4. Remove the capping solution by filtration and repeat **step 3** for 7 min.
5. Wash three times with DCM.
6. Check the disappearance of free amino groups by a convenient colorimetric test (*see* **Note 1**) and repeat the operation if necessary.

3.8. Difficult Sequences and Aggregation

When N^{α}-deprotection reactions and amino acid coupling steps do not go to completion or proceed in low yields, repeated or prolonged reaction times should overcome these problems. Nevertheless, this is not always sufficient. The cause of this failure is thought to be self-association of the peptide chain by hydrogen bond formation leading to aggregation. This aggregation results in incomplete

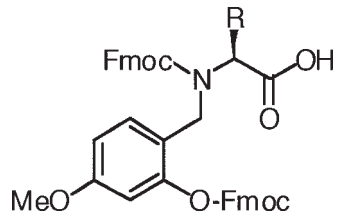


Fig. 8. Fmoc-(FmocHmb) amino acid.

solvation of the peptide-resin and inaccessibility of the reagents to the N-terminal amino group. This phenomenon is sequence-dependent with particular propensity for sequences containing high proportions of hydrophobic residues and can start from the fifth coupled residue. Two general methods can be used to disrupt the formation of secondary structures. The first consists of changing the peptide environment by adding chaotropic salts such as potassium thiocyanate or lithium chloride at a 0.4 *M* concentration, detergent solvents at 1% (v/v), or solvents such as DMSO, trifluoroethanol, or hexafluoroisopropanol to the reaction medium ***(22–28)***. The second approach involves structural changes in the peptide backbone itself by the introduction of protecting groups to selected amide bonds. This is usually achieved by using the 2-hydroxy-4-methoxybenzyl (Hmb) derivative of glycine. Other amino acids can also be incorporated as their *N*,*O*-*bis*(Fmoc)-*N*-(Hmb) derivatives approximately every 6–7 residues (**Fig. 8**) ***(29, 30)***. The incorporation of (Hmb) amino acids can be performed through their corresponding pentafluorophenyl esters, which are commercially available, or by standard coupling methods. For optimal *N*-acylation of the terminal Hmb residue that can be slow, it is recommended to use powerful coupling methods such as symmetrical anhydrides or preformed acid fluorides in DCM. The O-Fmoc protection of the Hmb derivatives is cleaved under piperidine treatment and the Hmb group in the final TFA cleavage. For Cys-, Ser-, and Thr-containing peptides their pseudoproline derivatives, also known to disrupt aggregation, can also be used ***(31–33)***.

3.9. Fmoc Deprotection

Removal of the temporary Fmoc protecting group from the N-terminus of the peptidyl-resin is normally achieved by short treatment with 20% piperidine in DMF (**Protocol 7**). The reaction is generally complete within 10 min, but can be longer in some cases. For safe removal a 20-min deprotection time is recommended. The deprotection results in formation of a dibenzofulvene-piperidine adduct that strongly absorbs in the UV range (*see* **Note 3**). Fmoc removal can be monitored by UV spectroscopy. This is a common procedure with automatic

synthesizers but not in standard manual SPPS. When incomplete deprotection is suspected, the use of 20% piperidine containing 1 to 5% DBU in DMF is recommended. However, DBU can promote aspartimide formation, thus its use should be avoided in Asp- or Asn-containing sequences.

Protocol 7. Removal of N^{α} Fmoc Protection

1. After coupling reaction and washings, resin is washed once with DMF.
2. Fill the SPPS reaction vessel (at least 2/3 of volume) with a 80:20 DMF/piperidine solution and stir for 20 min (*see* **Note 8**).
3. Remove the DMF/piperidine solution by filtration.
4. Wash the resin properly according to **Protocol 2**.

3.10. Final Cleavage

Concentrated trifluoroacetic acid is widely used for the simultaneous cleavage of the peptide from the resin and removal of side-chain protecting groups. In these conditions, side-chain protecting groups produce stabilized carbocations which can readily react with the electron rich side-chain of amino acids (Cys, Met, Tyr, Thr, Ser, Trp), leading to unwanted byproducts. To minimize this phenomenon, scavengers are added to the cleavage cocktail to trap carbocations. Numerous cocktails have been described to optimize the cleavage conditions of particular sequences. Silane-based cocktails that generally provide good results are nonodorous and less toxic than thiol-based cocktails ***(34)***. Special care should be taken to prepare and to use cleavage cocktails. All operations should proceed under a fume hood using gloves and safety glasses.

Protocol 8. Standard TFA Cleavage

1. Weigh the resin and place it in a round-bottom flask. When the resin is dry, swell it as described in **Protocol 1**.
2. Add 10 mL of cleavage cocktail (trifluoroacetic acid/water/triisopropyl silane 95:2.5:2.5) per 100 mg of resin.
3. Resin is cleaved for 90 min under gentle stirring (*see* **Note 9**).
4. Filter the resin on a glass frit and wash it twice with fresh cleavage cocktail. Recover the filtrate in a round-bottom flask.
5. Concentrate the cleavage cocktail *in vacuo* to approx 1/4 of its original volume.
6. Under vigorous stirring, add MTBE to precipitate the peptide. At least 10 times the initial TFA volume of MTBE should be added to precipitate the unprotected peptide. When the peptide does not precipitate, concentrate the solution in vacuo and go directly to **step 9**.
7. Filter the precipitate on a 4-porosity glass frit.
8. Triturate and wash by filtration the precipitated peptide on the frit three times with MTBE.

9. Solubilize the peptide in acetonitrile/water/TFA 50:50:0.1 and lyophilize. Solvent used for this step can be changed to increase solubility. Lyophilization can be repeated twice when scavengers are detected after reverse phase HPLC analysis of the peptide.

3.11. Disulfide Bridge Formation

One of the simplest ways to form disulfide bonds is oxidation of the fully deprotected linear peptide obtained after cleavage from the resin (in some cases, disulfide bridges began to form during the cleavage step) before or after preparative HPLC purification. The cyclization has to be performed using high dilution conditions to promote intramolecular disulfide bond formation versus intermolecular dimerization. Dimethyl sulfoxide is used to accelerate the oxidation reaction ***(35)***.

Protocol 9. Standard Formation of Disulfide Bridges by Air Oxidation

1. Dissolve the lyophilized peptide at 0.5 m*M* concentration in a 5% acetic acid solution.
2. Add 10% volume of DMSO in a large beaker. Adjust the pH to 6.0 to 7.0 if necessary with a 0.5 *M* ammonium acetate solution.
3. Stir vigorously the solution at room temperature for 12 h to incorporate atmospheric oxygen into the solution. Ideally, disulfide bridge formation should be monitored by reverse phase HPLC. Purification (**step 4**) should be performed immediately after the end of the reaction.
4. Purify the peptide by preparative reverse phase HPLC after acidification with a TFA solution (pH 2.0).
5. Lyophilize the solution containing the fractions.

For peptides having more than two cysteines, orthogonal protection of these residues can be used to allow selective and sequential disulfide bridge formation ***(5)***. However, especially if the peptide sequence is a natural one, simultaneous deprotection/bridge formation can be attempted in an adequate redox buffer. The equilibrium between the different disulfide bridge arrangements during "oxidative folding" leads, in most of cases, to the native folding.

Protocol 10. Formation of Disulfide Bridges
Using Oxidative Folding in Redox Buffer

1. Dissolve the linear peptide at a 0.05 to 0.1 m*M* concentration in a buffer (0.1–0.2 *M* Tris-HCl, pH 7.7–8.7) containing 1 m*M* EDTA and reduced (1–10 m*M*) and oxidized (0.1–1 m*M*) glutathione.
2. Stir the solution at 25 to 35°C and monitor the reaction by HPLC, normally from 16 to 48 h.
3. Lyophilize the solution after acidification with a TFA solution (pH 2.0).
4. Purify the oxidized peptide by preparative reverse phase HPLC.
5. Lyophilize the solution containing the fractions.

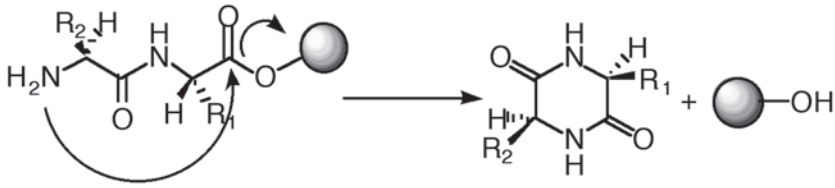


Fig. 9: DKP formation.

3.12. Side Reactions

Several residue- or sequence-dependent side reactions can occur during SPPS.

3.12.1. Diketopiperazine Formation

Diketopiperazine formation occurs at the C-terminal deprotected dipeptide stage by intramolecular cleavage of the resin ester linkage by the free amino function of the penultimate amino acid under basic conditions. This side reaction is particularly favored in Fmoc SPPS strategy as the base-induced deprotection of the Fmoc group releases a free amino group. The extent of its formation depends on the nature of the C-terminus amino acid and the type of peptide-linker ester anchor. Peptides containing proline or glycine in the C-terminal dipeptide sequence are especially sensitive to DKP formation. In such cases, hindered trityl-based resins should be used to avoid its formation and consequently the loss of the C-terminal dipeptide from the resin (**Fig. 9**).

3.12.2. Aspartimide Formation

The cyclization of aspartic acid residues to form aspartimide is the most likely side-reaction observed in routine SPPS (**Fig. 10**). This is a sequence-dependent side-reaction that occurs either during chain elongation or during final TFA cleavage when Asp(OtBu)-AA sequence (AA = Ala, Gly, Ser, Asn(Trt)) is present in the peptide. Hydrolysis of the aspartimide ring leads to a mixture of both α- and β-peptides. Its reaction with piperidine used for Fmoc removal also leads to the formation of α- and β-piperidides. Normally, in Fmoc-based SPPS, Asp (OtBu) provides sufficient protection. However, for particular sequences such as Asp(OtBu)-Asn(Trt) particularly sensitive to aspartimide formation, addition of HOBt to the piperidine solution or protection of the aspartyl amide bond with the 2-hydroxy-4-methoxybenzyl (Hmb) group should be considered ***(36)***.

4. Notes

1. **Kaiser test.** Solution A: 5% Ninhydrin in ethanol (w/v). Solution B: 80% phenol in ethanol (w/v). Solution C: KCN in pyridine (2 mL 0.001 *M* KCN in 98 mL pyridine). After washings, sample a few resin beads in a small glass tube and add 2 drops of each of the above solutions. Heat to 120°C for 4–6 min. Blue resin beads indicate the presence of resin-bound free amines.

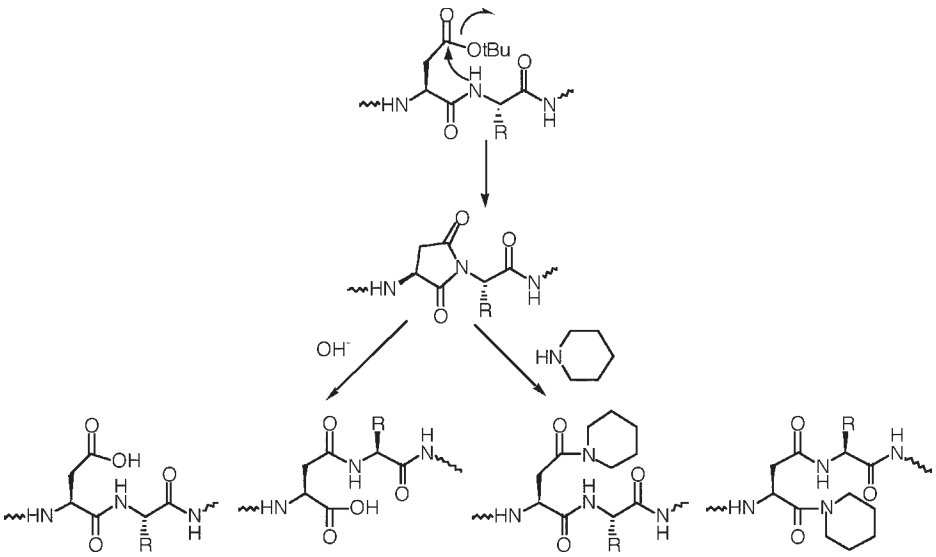


Fig. 10. Aspartimide formation and succinimide ring reopening in basic medium.

TNBS test. Solution A: 5% DIPEA in DMF (v/v). Solution B: 1% TNBS in water (w/v), commercially available from Fluka, St. Quentín-Fallavier, France 1% in DMF. Sample a few resin beads in a small glass tube, add 1 drop of each of the above solutions, and watch the sample under a microscope. The presence of resin-bound free amines is indicated within 10 min by yellow or red resin beads.

2. Polyamide-based and composite PEG-PS resin beads present a larger scope of solvent compatibility than PS beads, even allowing the use of water solutions. However, for routine peptide synthesis, solvents are the same than those used for PS-resin.
3. Fmoc determination: (1) Weigh duplicate samples of 5 to 10 mg loaded resin in an Eppendorf tube and add 1.0 mL 20% piperidine/DMF, stir for 20 min, and centrifuge down the resin. (2) Transfer 100 µL of the above solution to a tube containing 10 mL DMF and mix. (3) Set the spectrophotometer at 301 nm. Transfer 2 mL DMF into each of the two cells of the spectrophotometer (reference and sample cells), set the spectrophotometer to zero. Empty the sample cell, transfer 2 mL of the solution to measure, and check absorbance. (4) Substitution of the resin = [101 × (Absorbance)]/[7.8 × (weight in mg)] Check absorbance three times at 301 nm; calculate average substitution.
4. Reactivation of trityl resins: After swelling and washing the resin with toluene (five times), place the resin in a round-bottom flask and cover with toluene. Add freshly distilled acetyl chloride (1 mL/g resin) and fit a reflux condenser on the flask. Heat the reaction mixture at 60°C for 3 h. Allow to cool to room temperature, filter the resin on a frit, and wash six times with DCM. Load it immediately as described in **Protocol 4**.

5. To avoid the intramolecular cyclization of N-terminal Gln peptide leading to the formation of N-terminal pyroglutamic acid peptide, it is recommended to cleave the peptide from the support before removing the *N*-Fmoc terminal protection. The DMF/piperidine treatment has to be performed in solution followed by concentration of the mixture. The target peptide is then precipitated by MTBE and filtered.
6. To minimize epimerization during *N*-protected cysteine coupling, some coupling methods are recommended:
   a. Preferably use a hindered or weaker base such as 2,4,6-trimethylpyridine (TMP) or 2,6-di-tert-butyl-4-(dimethylamino)pyridine ***(37)***.
   b. Use 50:50 DMF/DCM coupling solvent instead of neat DMF.
   c. Avoid preactivation.
7. When a particularly acid-sensitive linker such as trityl is used, 50:50 DCM/acetic anhydride solution should be replaced by a 0.5 *M* solution of acetic anhydride and DIPEA in DMF. This will prevent premature cleavage of the linker by acetic acid released in the medium.
8. Some linkers, such as Rink amide linker, are sold *N*-Fmoc protected. The first removal of the Fmoc group may require a longer deprotection time particular with highly loaded resin (>0.9 mmol/g). In these cases it should be useful to repeat **steps 2** and **3** of **Protocol 7**.
9. Longer reaction times can be required to fully deprotect bulky side-chain protecting groups (Pbf, Trt…), especially when several protecting groups are close to each other in the sequence. When uncompleted deprotection is detected, the partially deprotected peptide should be cleaved in solution using larger cocktail volume and/or longer reaction time.

References

1. Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc*. **85,** 2149–2154.
2. Bergmann, M. and Zervas, L. (1928) Über katalytische racemisation von aminosäuren und peptiden. *Biochem. Z.* **203,** 280–292.
3. Goodman, M. and Levine, L. (1964) Peptide synthesis via active esters. IV. Racemization and ring-opening reactions of optically active oxazolones. *J. Am. Chem. Soc*. **86,** 2918–2922.
4. Carpino, L. A. and Han, G. Y. (1972) 9-Fluorenylmethoxycarbonyl amino-protecting group. *J. Org. Chem.* **37,** 3404–3409.
5. Lloyd-Williams, P., Giralt, E., and Albericio, F. (eds.) (1997) *Chemical Approaches to the Synthesis of Peptides and Proteins*. CRC Press, New York.
6. Czarnik, A. W. (1998) Solid-phase synthesis supports are like solvents. *Biotechnol. Bioeng. (Comb. Chem.)* **61,** 77–79.
7. Sherrington, D. C. (1998) Preparation, structure and morphology of polymer supports. *Chem. Commun.* 2275–2286.
8. Rink, H. (1987) Solid-phase synthesis of protected peptide fragments using a trialkoxy-diphenyl-methylester resin. *Tetrahedron Lett*. **28,** 3787–3790.

9. Bernatowicz, M. S., Daniels, S. B., and Köster, H. (1989) A comparison of acid labile linkage agents for the synthesis of peptide C-terminal amides. *Tetrahedron Lett.* **30,** 4645–4648.
10. Sieber, P. (1987) A new acid-labile anchor group for the solid-phase synthesis of C-terminal peptide amides by the Fmoc method. *Tetrahedron Lett*. **28,** 2107–2110.
11. Wang, S.-S. (1973) p-Alkoxybenzyl alcohol resin and p-alkoxybenzyloxycarbonylhydrazide resin for solid phase synthesis of protected peptide fragments. *J. Am. Chem. Soc*. **95,** 1328–1333.
12. Mergler, M., Nyfeler, R., Tanner, R., Gosteli, J., and Grogg, P. (1988) Peptide synthesis by a combination of solid-phase and solution methods II synthesis of fully protected peptide fragments on 2-methoxy-4-alkoxy-benzyl alcohol resin, *Tetrahedron Lett.* **29,** 4009–4012.
13. Flörsheimer, A. and Riniker, B. (1991) Solid-phase synthesis of peptides with the highly acid-sensitive HMPB linker, in *Peptides 1990: Proceedings of the 21st European Peptide Symposium* (Giralt, E. and Andreu, D. eds.), ESCOM, Lieden, pp. 131–133.
14. Barlos, K., Gatos, D., Kallitsis, J., et al. (1989) Darstellung geschützter peptidfragmente unter einsatz substituierter triphenylmethyl-harze. *Tetrahedron Lett*. **30,** 3943–3946.
15. Albericio, F. (2000) Orthogonal protecting groups for Nα-amino and C-terminal carboxylic functions in solid-phase peptide synthesis. *Biopolymers (Peptide Science)* **55,** 123–139.
16. Coste, J., Le Nguyen, D., and Castro, B. (1990) PyBOP®: A new peptide coupling reagent devoid of toxic by-product. *Tetrahedron Lett.* **31,** 205–208.
17. Knorr, R., Trzeciak, A., Bannwarth, W., and Gillessen, D. (1989) New coupling reagents in peptide chemistry. *Tetrahedron Lett*. **30,** 1927–1930.
18. Dourtouglou, V., Ziegler, J. C., and Gross, B. (1978) L'hexafluorophosphate de O-benzotriazolyl-*N*,*N*-tetramethyluronium: un réactif de couplage peptidique nouveau et efficace. *Tetrahedron Lett*. **19,** 1269–1272.
19. Carpino, L. A. (1993) 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* **115,** 4397–4398.
20. Carpino, L. A., El-Faham, A., Minor, C. A., and Albericio, F. (1994) Advantageous applications of azabenzotriazole (triazolopyridine)-based coupling reagents to solid-phase peptide synthesis *J. Chem. Soc. Chem. Commun.* **2,** 201–204.
21. Story, S. C. and Aldrich, J. V. (1994) Side-product formation during cyclization with HBTU on a solid support. *Int. J. Pept. Protein Res*. **43,** 292–296.
22. Westall, F. C. and Robinson, A. B. (1970) Solvent modification in Merrifield solid-phase peptide synthesis. *J. Org. Chem*. **35,** 2842–2844.
23. Yamashiro, D., Blake, J., and Li, C. H. (1976) The use of trifluoroethanol for improved coupling in solid-phase peptide synthesis. *Tetrahedron Lett.* **17,** 1469–1472.
24. Milton, S. C. and Milton, R. C. (1990) An improved solid-phase synthesis of a difficult-sequence peptide using hexafluoro-2-propanol. *Int. J. Pept. Protein Res.* **36,** 193–196.

25. Hendrix, J. C., Halverson, K. J., Jarrett, J. T., and Lansbury, P. T. (1990) A novel solvent system for solid-phase synthesis of protected peptides: the disaggregation of resin-bound antiparallel beta-sheet. *J. Org. Chem*. **55,** 4517–4518.
26. Hyde, C. B., Johnson, T., and Sheppard, R. C. (1992) Internal aggregation during solid phase peptide synthesis. Dimethyl sulfoxide as a powerful dissociating solvent. *J. Chem. Soc. Chem. Commun*. **21,** 1573–1575.
27. Stewart, J. M. and Klis, W. A. (1990) Peptides, polypeptides and oligonucleotides. Macro-organic reagents and catalysts and biomedical applications, in *Innovation and Perspectives in Solid Phase Synthesis and Related Technologies* (Epton, R., ed.), Mayflower Worldwide Ltd, Birmingham, UK, pp. 1–9.
28. Zhang, L., Goldhammer, C., Henkel, B., et al. (1994) Peptides, proteins and nucleic acids. Biological and biomedical applications, in *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Mayflower Worldwide Ltd, Birmingham, UK, pp. 711–716.
29. Bedford, J., Hyde, C., Johnson, T., et al. (1992) Amino acid structure and "difficult sequences" in solid phase peptide synthesis. *Int. J. Pept. Protein Res.* **40,** 300–307.
30. Hyde, C., Johnson, T., Owen, D., Quibell, M., and Sheppard, R. C. (1994) Some difficult sequences made easy. A study of interchain association in solid-phase peptide synthesis. *Int. J. Pept. Protein Res.* **43,** 431–440.
31. Mutter, M., Nefzi, A., Sato, T., Sun, X., Wahl, F., and Wohr, T. (1995) Pseudo-prolines (psi Pro) for accessing "inaccessible" peptides. *Peptide Res.* **8,** 145–153.
32. Wohr, T., Wahl, F., Nefzi, A., et al. (1996) Pseudo-prolines as a solubilizing, structure-disrupting protection technique in peptide synthesis. *J. Am. Chem. Soc*. **118,** 9218–9227.
33. Guichou, J. F., Patiny, L., and Mutter, M. (2002) Pseudo-prolines (Pro): direct insertion of Pro systems into cysteine containing peptides. *Tetrahedron Lett*. **43,** 4389–4390.
34. Pearson, D. A., Blanchette, M., Baker, M. L., and Guindon, C. A. (1989) Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis. *Tetrahedron Lett*. **30,** 2739–2742.
35. Tam, J. P., Wu, C. R., Liu, W., and Zhang, J. W. (1991) Disulfide bond formation in peptides by dimethyl sulfoxide: scope and applications. *J. Am. Chem. Soc*. **113,** 6657–6662.
36. Quibell, M., Owen, D., Packman, L. C., and Johnson, T. (1994) Suppression of piperidine-mediated side product formation for Asp(OBut)-containing peptides by the use of N-(2-hydroxy-4-methoxybenzyl)(Hmb) backbone amide protection. *J. Chem. Soc. Chem. Commun*. **20,** 2343–2344.
37. Han, Y., Albericio, F., and Barany, G. (1997) Occurrence and minimization of cysteine racemization during stepwise solid-phase peptide synthesis. *J. Org. Chem*. **62,** 4307–4312.

2

Chimerism

A Strategy to Expand the Utility and Applications of Peptides

John Howl

Summary

The modular nature of peptides can be exploited in the synthesis of chimeric sequences that combine diverse motifs in a single molecule. A theoretical consideration of the classification of peptides further expounds the multigeneric nature of peptide chimeras. Strategies for chimeric peptide syntheses include the chemical cross-linking of monomers and tandem combination by conventional SPPS. Additional details of chimeric peptide synthesis are also provided elsewhere in this volume. This chapter also explores some of the more common applications of chimeric peptides with particular emphasis on the molecular pharmacology of sequences that include address motifs for G protein-coupled receptors. Specific details of the biological properties of chimeras containing mastoparan, an amphiphilic tetradecapeptide component of wasp venom, further illustrate the novel and often unpredictable biological actions of chimeric constructs. These and numerous additional studies confirm that chimerism is an established strategy for the synthesis of molecular probes and bioactive agents.

Key Words: Bradykinin; chimerism; galparan; G protein-coupled receptor; ligand binding; mastoparan; secretion; vasopressin.

1. Introduction

The continual and necessary refinement of both synthetic methodologies and materials has clearly established peptide synthesis as a common discipline at the interface of chemistry and biology. Predictably, the range of applications in which peptides have demonstrated utility has also rapidly increased in recent years. The Peptide Institute, Protein Research Foundation (PRF), Osaka (www.

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

prf.or.jp) currently lists 233,109 entries (83,164,416 residues) in a database of real and predicted peptide and protein sequences (SEQDB). A similar number of entries (298,109) are also to be found in the database of synthetic compounds that includes unnatural amino acids and chemically synthesized peptides (SYNDB). Though biologically active peptides usually comprise fewer than 50 residues, mostly coded or proteinogenic amino acids, the PRF revised its definition of the term "peptide" in 1998 to also include sequences of 50 to 100 residues.

The size differential between a large peptide of 100 residues and a small protein, or perhaps a protein domain, is a purely arbitrary distinction. However, both (1) the technology-driven acceleration in the rate of new peptide generation and (2) the increased size limit from 50 to 100 residues will vastly increase the number of sequences that need to be accommodated within the PRF databases. The problem is exacerbated by the fact that there is a temporal continuum to peptide production. Thus, it is impossible to know how many *bona fide* peptides are missing for the existing databases. Many of these sequences will have been erroneously produced and immediately discarded or simply never reported in the scientific literature. Some of these unaccounted-for peptides have no doubt been physically lost and now exist only in a virtual sense. Should peptide databases include ghosts? Unfortunately, such questions are relatively trivial when one considers future implications. The total number of all possible peptides must be finite, but what a number! Moreover, it is entirely reasonable to speculate that the sequences contained within the two current PRF databases are but a tiny fraction of this total.

Databases are most useful if they can be interrogated to extract information and provide answers to simple questions. These processes might themselves promote a logical subdivision of the database, a reactive classification of entries. Alternatively, some form of comparative analysis could be used to provide order to the database, a more proactive process of classification. The question arises as to how might all peptide sequences be arranged to provide some degree of logical order. One possibility would be to compare only primary sequences, a primitive though exhaustive form of parsimony analysis. Such an approach would assemble a few clusters of sequence-related peptides, such as opioid receptor ligands. The problem is, however, that sequence analysis would not identify most groups of peptides with similar biological roles. The diversity of sequences pertaining to peptide hormones and neuropeptides is testimony to this point. Thus, a meaningful subdivision of the database of all peptide sequences would require the use of a complex algorithm that compares sequence, structure, and biological activities.

The next question to arise from contemplation of the database of all peptide sequences is how best to represent relationships between different sequences.

Two dimensions are clearly insufficient to describe all possible relationships among all possible sequences. One approach to this problem is to imagine all peptide sequences existing in a vast space through which it is possible to travel and explore their interrelationships; the hypothetical space of all possible peptide sequences (peptide space). A similar concept of a hypothetical space containing gene kits was eloquently developed as a model to explore genetic relationships and the origins of life ***(1)***, and the following description of peptide space owes much to this work. The borders around peptide space identify it as a subdivision of organic chemical space, a larger volume containing all possible carbon-based chemical entities. Within peptide space one might imagine each hypothetical peptide sequence to exist within a stackable box that contains a small-scale homogeneous reference sample and documentation providing contemporary synthetic instructions and appropriate analytical data (m.p., AA analysis, mass spectra, etc.). Moreover, the number of opened boxes colored green, representing those sequences that have been synthesized and might be found in the SEQDB and SYNDB databases, would be mere specks of color in a 3D (or 4D) spatial continuum of unopened red boxes. We might also assume that peptide space is organized by the complex algorithm required to take account of sequences, structure, and biological roles. By analogy to the even larger space of all possible molecules, in which we find clusters of opened boxes containing broadly similar polymeric structures (genes, proteins, lipids, plastics, rock crystals, etc), the distribution of green boxes in peptide space is certainly not random.

As we travel through peptide space we observe clusters of green boxes indicative of a higher order of organization. This spatial organization of peptides into representative groups, a kind of retro-molecular-cladism, is a physical map of the different functions and/or applications of mostly extant peptides. Depending on the bias of the sorting algorithm, we might logically expect to find clusters of green boxes that could include groups of peptides such as hormones, neuropeptides, enzyme inhibitors, and secondary structure mimetics. The spatial relationships between boxes will of course require constant modifications and fine-tuning as more data become available to influence the sorting algorithm. It is intriguing to consider how Merrifield's introduction of solid phase synthesis, and significant modifications including the Fmoc strategy (*see* Chapter 1), has accelerated the rate of box opening.

One additional important concept to consider is that of peptide classification. If each sequence in peptide space is considered an individual *species*, then our clusters of green boxes are representative of a higher order of classification, the *genus*. With this concept in mind, **Table 1** is an attempt to produce a *generic* classification of peptides. The number of individual species in each genus is, of course, variable and no doubt the future will see the *evolution* of new genera

Table 1
A Possible Classification Scheme of Peptides

Class. biologically active sequences
Order. endogenous mediators
Hormone
Neuropeptide
Pheromone
Toxin
Protein fragment
Natriuretic
Cytokine
Chemokine
Growth factor
Order. antagonists and inhibitors
Receptor antagonist
Integrin antagonist
Enzyme inhibitor
Antimicrobial
Antitumoral
Antibiotic
Order. protein modulator
Ion channel modulator
Enzyme modulator
Receptor modulator
Modulator of protein–protein interaction
Molecular switch
Secretagogue

and higher orders recognizable as green clusters and their interrelationships in the red continuum. An algorithm that compared only primary sequences would produce an entirely different classification; some peptides have multiple biological actions allied to multiple genera.

One might also imagine the dramatic increase in volume of the space of all possible peptides required to accommodate the PRF's redefinition and expansion of qualifying sequence length. On the contrary most, if not all, chimeric peptides studied to date already existed in the pre-1998 space, but we have only yet opened a very small fraction of the total number of dusty red boxes containing this intriguing class of biomolecule. I use the term "class" here in a strictly nonhierarchical, nonbiological sense as the majority of chimeric pep-

Table 1 (Continued)

Order. cellular delivery system or component
Cell-specific homing sequence
Organelle targeting sequence
Signal sequence
Cell-penetrating vector
Delivery vehicle
Order. immunological peptide
Antigen
Immune cell active sequence
Chemotactic
Immunosuppresssant
Antibody fragment
Class. nonbiologically active sequences
Order. carrier peptide
C-peptide
Neurophysin
Preprohormone
Prohormone
Order. structural mimetic
Secondary structure mimetic
Protein domain mimetic
Enzyme substrate
Pore former
Order. molecular support
Linker
Scaffold
Regioselectively addressable functionalized template

This scheme is almost certainly incomplete and some genera, including receptor antagonists, can be further divided using a variety of classification criteria. There are also many other peptides generated by the enzymatic processing of larger precursors that might be included in the order of carrier peptides.

tides synthesized and studied to date are essentially hybrid species comprising recognizable sequences from two or more ancestors, often from different genera. A relatively simple algorithm that considers only sequences would have little organizational exasperation with chimeras. However, the complex algorithm that organized peptide space using sequence, structural, and functional data might have labored to accommodate single boxes containing chimeric peptides. As we shall see, chimeric peptides can display multiple biological actions

that may or may not reflect their ancestral components. Thus, peptide space might need to include multiple clones of some boxes to accommodate the multigeneric characteristics of their contents.

One purpose of this chapter is to highlight some key findings related to the properties and applications of chimeric peptides. No doubt the future will see a tangible change in the balance and distribution of green and red boxes in the subspace of all possible chimeric peptides.

1.1. Development of Chimeric Ligands for G Protein-Coupled Receptors

G Protein-coupled receptors (GPCRs) are the generic molecular target for approx 50% of all current drugs and a very common target for peptide mediators that include hormones and neuropeptides. Moreover, we can predict that more peptide mediators await discovery and that some of these will prove to be the ligands for the numerous orphan GPCR gene sequences identified by molecular cloning and *in silico* analyses of draft genome sequences. The differential expression of GPCRs in central and peripheral sites enables drug action to be selectively directed to appropriate cells and tissues.

The small nonapeptide hormones [Arg^8]vasopressin (AVP; H-CYFQNCPRG-NH_2; intramolecular bond between Cys residues) and bradykinin (BK; H-RPP GFSPFR-OH) were among the first biologically active peptide hormones to be isolated some four decades ago, and are now known to collectively modulate most physiological and many pathological processes. Extensive characterization of numerous structural analogs of AVP and BK has since provided extensive knowledge of the SAR enabling both GPCR binding and activation ***(2,3)***. The molecular cloning of cDNAs encoding a variety of AVP and BK receptor subtypes confirmed that these proteins display the common 7-transmembrane architecture of rhodopsin-like (Type I) GPCRs ***(4,5)***.

The overwhelming majority of studies that defined, and continue to define, the SAR of peptidyl analogs of both AVP and BK have concentrated on the synthesis of individual nonapeptides or shorter deletion analogs ***(2,3)***. However, amino-terminal extended analogs of both hormones, including Val-Asp-AVP ***(6)*** and Lys^0-BK (kallidin) ***(4)***, are also biologically active. A report in 1995 ***(7)*** also indicated that sequences derived from BK (residues 2–9 and 2–8) could be utilized in the design of galanin receptor antagonists as a carboxyl-terminal extension of galanin (1–13). (Indeed, my personal interest in chimeric GPCR ligands was fuelled by hearing Ülo Langel describe these and other studies at around the same time.) Intriguingly, the same galanin-derived sequence as an amino-terminal extension of mastoparan (MP), which has been labeled galparan, is a galanin receptor ligand and a potent secretagog ***(8)***. Observations

Table 2
Some Pharmacological Properties of Chimeric Ligands for GPCRS

Peptide Sequence	Comments
1. AVP-BK H-CYFQNCPRGRPPGFSPFR-OH	K_d B_2 bradykinin receptor = 52 n*M*. 20% Antidiuretic activity of AVP; 20–30% activity of oxytocin in uterotonic tests; Hypotonic effect comparable to BK ***(35)***.
2. H22 [Phaa$_D$Tyr(Et)^{2}Arg^6]AVP(1-8)-εAhx-$_D$Arg^0[Hyp^3$_D$Phe^{7}Leu^8]BK(1-9)	Combines a V_{1a}-selective vasopressin antagonist and a B_{2a}-selective antagonist joined by a flexible εAhx linker. High affinities for both V_{1a} (K_d = 0.3 n*M*) and B_2 (K_d = 4.8 n*M*) receptors.
3. H28 [Phaa$_D$Tyr(Et)^{2}Arg^6]AVP(1-9)-DF-NH_2	A V_{1a}-selective antagonist with a carboxyl-terminal extension providing the integrin anatgonist sequence *RGDF*. Retains high affinity (K_d = 0.4 n*M*) for the V_{1a} vasopressin receptor. Antagonizes both uterotonic (PA_2 = 7.8) and pressor (PA_2 = 7.2) activities of oxytocin and vasopressin respectively ***(35)***.

AVP-BK displays negligible affinity for neurohypophysial peptide receptors in binding assays (*see* **Subheading 3.2.**) but is an agonist at the B_2 receptor to which it binds with moderate affinity. However, bioassays indicate that AVP-BK possesses both antidiuretic and uterotonic activities. These observations might indicate some form of enzymatic processing to generate a smaller biologically active fragment. H22 and H28, high-affinity antagonists for the V_{1a} vasopressin receptor, present sequences of a B_2 receptor antagonist and a fibronectin-related integrin binding sequence respectively.

$_D$Tyr(Et) is *O*-ethyl-D-tyrosine; εAhx is aminohexanoate; Phaa is phenyacetyl; Hyp is hydroxyproline.

such as these confirm the multigeneric character of chimeric peptides. Moreover, studies with similar peptides also lead to the discovery of the cell penetrant transportans (*see* Chapter 5).

A first report detailing some biological properties of chimeric constructs combining AVP and BK appeared in 1997 ***(9)***. The broad aim of these initial studies was to determine whether single peptides combining address sequences for AVP and BK receptors retained binding affinities. These and a variety of subsequent studies have clearly revealed that chimeric peptides can selectively bind GPCRs for both AVP and BK. **Table 2** reveals the structures and some biological properties of chimeric GPCR ligands that bind receptors for AVP and BK.

1.2. M375 and M391: Case Study to Illustrate Multigeneric Character of Chimeric Peptides

Additional details of these two peptides illustrate both the unique properties of the MP sequence and the fact that chimeric peptides interact with multiple protein targets.

M375 [Phaa$_D$Tyr(Me)^{2}Arg^{6}Tyr^9]AVP(1-9)-MP
M391 [Phaa$_D$Tyr(Me)^{2}Arg^{6}Tyr^9]AVP(1-9)-εAhx-MP

These are primary sequences of the chimeric peptides M375 and M391. $_D$Tyr (Me) is *O*-methyl-D-tyrosine; Phaa is phenylacetyl; εAhx is aminohexanoate.

M375 and M391 were synthesized in Stockholm as part of a program to evaluate the biological properties of peptides combining sequences of GPCR receptor ligands with MP. [Phaa$_D$Tyr(Me)^{2}Arg^{6}Tyr^9]AVP, a linear antagonist of AVP, selectively binds the commonly expressed V_{1a} subtype of vasopressin receptor ***(10)***. M375 (K_d = 3.8 n*M*) and M391 (K_d = 9.0 n*M*) also display selective, high-affinity binding to the V_{1a} receptor ***(10)***. These data clearly indicate that substantial carboxyl-terminal extension is compatible with the binding of antagonists to the V_{1a} receptor protein. However, these peptides also interact with at least three other unrelated proteins. Both M375 and M391 are potent insulin secretagogs acting at a distal site in the secretory pathway distinct from that modulated by MP alone ***(11)***. These observations indicate that MP-containing chimeras translocate across the plasma membrane of cells to specifically interact with intracellular proteins that probably include heterotrimeric G proteins. This interesting property of MP chimeras has since been exploited in the development of transportan and its deletion analogs as a novel class of synthetic cell-penetrating peptide. M375 and M391 also interact with proteins involved in intracellular calcium homeostasis. Both peptides inhibit ER and SR Ca^{2+}-ATPases and promote calcium release by activating the ryanodine receptor in vitro ***(12)***.

1.3. Other Chimeric Peptidyl Secretagogs

The observations that M375 and M391 are novel insulin secretagogs prompted more recent studies to engineer additional chimeric MP analogs that promote mast cell secretion ***(13)***. These studies identified differential peptidyl secretagogs (DPS) that selectively stimulate the exocytosis of 5-HT or the secretory granule marker β-hexosaminidase respectively. Moreover, many of these MP analogs are significantly more potent secretagogs than MP itself. Our recent contributions to this area have focused on chimeric peptides combining MP with either sequences related to small peptide hormones and neuropeptides or specific peptides known to modulate secretion. Of more than 70 peptides we have screened, MP S (INWKGIASM*Aib*RQVL-NH_2), a synthetic MP analog con-

taining α-aminoisobutyric acid (Aib), a known helix promoter, is the most potent and selective activator of β-hexosaminidase secretion. In contrast, M436 (INL KALAALAKKILVTHRLAGLLSRVPTNVGSKAF-NH_2), a chimeric construct combining MP with sequences of human calcitonin gene-related peptide, is the most selective and potent activator of 5-HT secretion from RBL-2H3 ***(13)***. Intriguingly, M436 is also an antagonist of the $CGRP_1$-like receptor ***(14)***, further indicating that chimeric peptides can display both receptor- and non-receptor-mediated activities.

1.4. Additional Applications of Chimeric Peptides

Predictably, peptide chimeras have found utility for many additional applications other than those listed here. A detailed discussion of the growing literature pertaining to peptide chimeras, particularly GPCR ligands, is beyond the scope of both this chapter and this book. However, to partly address my own bias toward the study of peptide hormones and secretagogs, brief mention is made of a few notable examples of chimeric peptides using sequences derived from other peptide genera. The sequence and some properties of vasonatrin, a 27-AA chimera of atrial natriuretic peptide (ANP) and C-type natriuretic peptide (CNP), were reported in 1993 ***(15)***. Intriguingly, in vivo studies indicated that vasonatrin possessed venodilatory actions (CNP), natriuretic actions (ANP), and a unique vasodilatory action in arteries characteristic of neither CNP nor ANP ***(15)***. A substance P-opioid chimeric peptide (YPFFGLM-NH_2) has been reported as a prototypic, ant-tolerance-forming analgesic ***(16)***. These unique properties of YPFFGLM-NH_2 appear to be the consequence of the co-activation of both μ-opioid receptors and substance P receptors in the spinal cord. Chimeric analogs of other neuropeptides that include opioids, deltorphins, Met-enkephalin, and FMRFa have also been developed as novel receptor probes ***(17,18)***.

There is no doubt that the range of peptide genera included in chimeric design will increase. One important development is the application of homing sequences to deliver peptides, drugs, or nanoparticles to tumor cells or neovasculature ***(19–23)***. Such sequences include (1) CGFECVRQCPERC, which binds lung endothelial cell membrane dipeptidase ***(19)***; (2) KDEPQRRSARLSAKPAPPKPEP KPKKAPAKK, which binds cells and blood vessels in selected tumours ***(20)***; (3) CGNKRTRGC, which recognizes tumor cells in lymphatic vessels ***(21)***; and (4) SMSIARL, which binds to vasculature in the human prostate gland ***(22)***. These sequences, together with a range of GPCR ligands as reviewed in **ref. *24***, offer enormous potential for the selective delivery of cytotoxic agents to tumor cells and neovasculature.

One final development worthy of note here is the development of novel anticancer agents that are dimeric bradykinin analogs. There has been a long-term interest in cross-linked kinin analogs as both agonists and antagonists of GPCRs

(9,25). However, more recent studies have indicated that dimeric bradykinin antagonists such as CU210 specifically inhibit the growth of lung cancer lines ***(26)***. Such findings are particularly interesting since monomeric bradykinin agonists and antagonists lack this activity.

It is entirely reasonable to suggest that other chimeric peptides have intrinsic biological properties that await discovery through rational analyses. Increased understanding of these properties might provide generic guidelines for chimeric peptide construction. Moreover, as more data become available, peptide space will require constant reorganization to assimilate this new information. As the number of green boxes increases, their spatial organization may well indicate new peptide genera or, indeed, traces of some other hierarchical classification. However, it is very difficult to envisage when, where, and how such a process might end.

We know nothing of the practical or conceptual limitations of chimeric peptide design and two potential caveats are apparent.

First, how should the algorithm that organizes peptide space deal with failures? The chimeric peptide vasokinin, BK(1)-AVP(2-5)-BK(6-9) (H-RYFQN SPFR-OH), combines motifs from the amino and carboxyl termini of BK with a middle section from AVP. The peptide was conceived as a putative GPCR ligand but lacks affinity for receptors that bind BK or AVP ***(27)***. However, vasokinin is a competitive inhibitor of angiotensin-converting enzyme and could well display other biological activities that remain to be discovered. Where might vasokinin be located in the subspace of chimeric peptides, and should there be multiple boxes?

Second, it is most likely that the peptide classification scheme presented here will soon be outdated. Indeed, any such scheme is largely subjective and I apologize here for any notable omissions or inconsistencies. Future developments will no doubt require an increasingly complex classification system. Imagine the futility of attempting a rational classification of an animal kingdom that contained organisms with every possible body plan ranging from unicellular to multicllular, from diploblastic to triplobastic, and beyond. A multitude of reproductive strategies would ensure that our all-encompassing animal kingdom would collectively occupy every conceivable niche the home planet could offer. Thus, in a sense it is the missing animals that enable a logical and hierarchical classification of this planet's fauna. Evolutionary pressures will no doubt influence the opening of boxes in peptide space with successful sequences and new discoveries serving as templates for further modification and improvement. Certain clusters of green boxes will, therefore, rapidly expand in all dimensions at the expense of other domains that remain largely unexplored. However, we may reach a time when the sheer abundance of green boxes obscures any obvious spatial organization either of peptide space or its subspace of chimeric

peptides. Perhaps when the system becomes this complex it would be better to abandon any organizational algorithm using functional data and classify all peptides using primary sequence only. At this point in time the concept of chimerism would essentially become redundant and a single green box containing vasokinin would be located near other sequence-related nonameric peptides with quite different biological properties.

2. Materials

A generic list of materials pertinent to the detailed evaluation of chimeric peptide is provided.

1. Chimeric peptide(s).
2. Appropriate assay system components.

Clearly, a detailed list of all materials necessary to determine the many activities of chimeric peptides is not possible here. As this article has focused on chimeras that combine GPCR ligands with MP it is appropriate to include details of two assay systems that can be used to determine both receptor binding affinity and secretory efficacy. The details pertaining to receptor binding analyses relate specifically to studies with the vasopressin V_{1a} receptor ***(5)***, though they are broadly applicable to all other GPCRs that bind peptide ligands. The rat basophilic line RBL-2H3 is a convenient mast cell model for studying peptide-induced secretion of β-hexoseaminidase ***(13)***.

2.1. V_{1a} Receptor Binding Analysis

1. Rodent liver membranes (*see* **Note 1**).
2. Binding assay buffer: 20 m*M* HEPES, 10 m*M* $Mg(CH_3COO)_2$, 1 m*M* EGTA, 1 mg/mL BSA, pH 7.4.
3. Tritiated tracer ligand: [^{3}H]AVP (agonist) or [^{3}H][d$(CH_2)_5$[Tyr(Me)^2]AVP.
4. (V_{1a}-selective antagonist ***[5]***) (*see* **Note 2**).
5. [d$(CH_2)_5$[Tyr(Me)^2]AVP (*see* **Note 2**).
6. Chimeric peptide(s).
7. Circulating water bath.
8. High-speed bench centrifuge.
9. Tissue solubilizer (Soluene 350).
10. Scintillation fluid.
11. Liquid scintillation counter.

2.2. β-Hexoseaminidase Secretion From RBL-2H3

1. RBL-2H3 cells in Dulbecco's modified Eagle's medium (DMEM).
2. Chimeric peptides.
3. Balanced salt solution (BSS).
4. 0.1% v/v Triton X-100.
5. 96-Well plates.

6. 1 m*M* ρ-nitrophenyl *N*-acetyl-β-D-glucosamide in 0.1 *M* sodium citrate pH 4.5.
7. 0.1 *M* Na_2CO_3/$NaHCO_3$, pH 10.5.
8. Microplate reader at 405 nm.

3. Methods

The provided methods describe (1) the design and synthesis of peptide chimeras, (2) the analysis of receptor binding affinity, and (3) β-hexoseaminidase secretion from RBL-2H3. These protocols have proven particularly useful for the development of both GPCR ligands ***(9)*** and differential secretagogs ***(13)***.

3.1. Synthesis of Chimeric Peptides

Chimeras synthesized by tandem linkage (*see* **Subheading 3.1.1.**) is the quantitatively dominant form of construct studied to date. The relative ease of synthesis using conventional methodologies (*see* Chapter 1) is one explanation for the relative abundance of tandem constructs. A brief discussion of homo- and hetero-dimeric peptides constructed by the covalent crosslinking of monomers is also provided, in **Subheading 3.1.2.**

3.1.1. Tandem Chimeric Constructs

M375 and M391 are examples of tandem constructs synthesized according to the following generic template:

(D)-[SEQUENCE A]-[LINKER C]-[SEQUENCE B]-(E)

This is a simple generic template for chimeric peptide construction. For the design of chimeric GPCR ligands, both address and message motifs have been used as sequences **A** and **B**. The physical linkage (**C**) joining sequences **A** and **B** is usually one of three chemical entities: (1) a simple peptide bond (e.g., M375); (2) a flexible aminohexanoic acid spacer (e.g., M391) designed to release the constraint of a relatively rigid peptide backbone; or (3) another amino acid or sequence of amino acids. Examples of the latter strategy include Gly-Gly, used to join the cytotoxic sequence $_D$(KLAKLAK)$_2$ to a prostate homing sequence ***(22)***; and lysine, which provides a primary amine for the attachment of other moieties. A variety of common modifications at the amino terminal (**D**) and the carboxyl terminal (**E**) are feasible with conventional SPPS. Thus, it is clear that even the simple template indicated here offers enormous potential for generating a vast array of structural chimeras.

3.1.2. Cross-Linked Dimeric Peptides

The covalent linkage of peptide monomers is an alternative strategy for the synthesis of peptide chimeras. The formation of cystine via the reduction of paired Cys residues is perhaps the most commonly employed methodology for

achieving this strategy. Cystine formation can be achieved by simply mixing free thiol-containing peptides in weakly alkaline conditions. However, such a strategy is compatible mostly with the formation of homodimers, as multiple dimers are produced when mixing different thiol-containing monomers. The 3-nitro-2-pyridinesulfenyl (Npys) protecting group, first introduced by Bernatowicz et al. ***(31)***, is a particularly useful side-chain protecting group for Cys that provides an active leading group to facilitate the formation of unsymmetrical disulfide bonds by Npys-thiol exchange. Such a strategy can be used to conjugate thiol-containing cargoes to transportans and other cell penetrating peptides. Further details of the use Cys (Npys) are provided in Chapter 5. The vast array of other strategies employed for the synthesis of chimeras from peptide monomers is beyond the scope of this chapter. Bifunctional cross-linkers can be used to join unprotected peptides together, while Tam's group ***(32)*** has developed a numerous chemoselective methodologies. The latter approaches, based largely on thiol and carbonyl chemistries ***(32)***, offer enormous potential for the covalent coupling of peptide monomers and are perhaps not as commonly employed as they deserve to be.

3.2. V_{1a} Receptor Binding Analysis

Competitive displacement of a radiolabeled tracer ligand (*see* **Note 2**) is the usual method to determine the receptor binding affinity of a GPCR ligand. Additional details of these methods are provided elsewhere ***(28)***.

1. Thaw rat liver membranes on ice and dilute in binding assay buffer to a final protein concentration of 0.208 mg/mL.
2. Add 10 µL tracer ligand and 10 µL competing peptide to membranes to give a final volume of 500 µL, containing 100 µg of membrane protein. Using this system a concentration of tracer ligand from 0.2–1.0 × 10^{-9} *M* provides a good specific: nonspecific binding ratio. Unlabeled peptides are usually added at a range of concentrations from 10^{-11} *M* to 10^{-6} *M*.
3. Incubate samples in 1.5-mL Eppendorf tubes at 30°C in a circulating water bath. Equilibrium binding is achieved after 90 min incubation.
4. To separate bound and free radioligand, sediment membranes by centrifugation (5 min, 13,000*g*) and wash pellets twice.
5. Dry Eppendorf tubes and solubilize membranes with 50 µL tissue solubilizer.
6. Add scintillation fluid and determine radioactivity by liquid scintillation spectroscopy.

K_d values of unlabeled peptides can be calculated from experimentally derived IC_{50} values by correcting for ligand occupancy of the tracer ligand according to the equation:

$$K_d = IC_{50} \times [K_d^*/(L^* + K_d^*)]$$

where K_d^* is the dissociation binding constant of the tracer ligand and L* is the free tracer ligand concentration. L* can be accurately determined by counting a 100-µL aliquot of the supernatant after membrane sedimentation. A 1000-fold excess of a suitable high affinity antagonist (*see* **Note 2**) is usually included in the assay to define the nonspecific binding component of the tracer ligand.

AVP and its analogs are relatively stable when incubated with rat liver membranes; there is little or no degradation of peptides during the 90 min incubation required to ensure binding has reached equilibrium. However, other peptide ligands including BK are more labile in the presence of biological membranes and may require the inclusion of a cocktail of protease inhibitors to prevent ligand, and possibly receptor, degradation ***(29)***. It is also noteworthy that some chimeric GPCR ligands, including galanin analogs ***(33)***, bind with Hill coefficients significantly <1 indicating complex receptor interactions with more than one recognition site.

3.3. β-Hexoseaminidase Secretion From RBL-2H3

RBL-2H3 is a widely used model of mucosal mast cells ***(34)*** employed to study both constitutive and regulated secretory pathways. Significantly, RBL-2H3 does not respond to a range of polybasic peptides that promote secretion from other mast cells but is sensitive to MP and its analogs ***(13)***. The secretion of β-hexoseaminidase, a secretory lysosomal marker, is a convenient assay of the efficacy of peptidyl secretagogs that correlates with the activation of phospholipase D, enzyme intimately involved in regulated secretion ***(13)***. Secretion assays are generally performed in a 24-well plate format using subconfluent cell monolayers.

1. Add peptidyl secretagogs to cells for a period of 15 min at 37°C in a final volume of 0.25 mL.
2. Collect medium and remove cell debris by centrifugation.
3. Secreted β-hexoseaminidase is assayed in samples of cell medium. Cellular β-hexoseaminidase activity is determined by washing cells in BSS and adding 100 µL of 01.% v/v Triton X-100 to extract residual β-hexoseaminidase from cell monolayers.
4. Transfer 5-µL samples of medium or cellular extract to 96-well plates and incubate with 20 µL of 1 m*M* ρ-nitrophenyl *N*-acetyl-β-D-glucosamide in 0.1 *M* sodium citrate pH 4.5, for 1 h at 37°C.
5. Add 200 µL of 0.1 *M* $Na_2CO_3/NaHCO_3$, pH 10.5, and determine β-hexoseaminidase activity by colorimetric analysis on a microtiter plate reader at 405 nm.

4. Notes

1. Rat liver is the most convenient source of tissue for the preparation of crude plasma membranes that contain the V_{1a} subtype of vasopressin receptor. Methods for the

production of rat liver membranes that can be conveniently stored at –20°C are presented elsewhere ***(28)***. Detailed protocols for the production of kidney medulla membranes of rodent and bovine origin are also available and such preparations are suitable for studying binding to V_2 vasopressin and B_{2a} bradykinin receptors ***(28,29)***.

2. Both [^{3}H]AVP and [^{3}H][d(CH_2)$_5$[Tyr(Me)^2]AVP, a V_{1a}-selective antagonist, are suitable radiolabeled tracer ligands for the V_{1a} receptor. A high concentration of [d(CH_2)$_5$[Tyr(Me)^2]AVP or another analog can be used to define the nonspecific binding of the radioligand.

Acknowledgments

A variety of sources have provided financial support for these studies, including grants from both the Wellcome Trust and the Royal Society. A number of researchers, past and present, have contributed to these investigations and the author is particularly indebted to the efforts of both Michelle Farquhar and Sarah Jones.

References

1. Woolfson, A. (2000) *Life Without Genes.* Harper Collins, London, U.K.
2. Manning, M., Bankowski, K., and Sawyer W. H. (1987) Selective agonists and antagonists of vasopressin, in *Vasopressin* (Gash, D. M. and Boer, G. J., eds.), Plenum Press, New York, pp. 335–368.
3. Stewart, J. M. (1995) Bradykinin antagonists: developments and applications. *Biopolymers* **37,** 143–155.
4. Howl, J. and Payne, S. J. (2003) Bradykinin receptors as a therapeutic target. *Expert Opin. Ther. Targets* **7,** 277–285.
5. Howl, J. and Wheatley, M. (1995) Molecular pharmacology of V_{1a} vasopressin receptors. *Gen. Pharmacol.* **26,** 1143–1152.
6. Gitelman, H. J., Kalpper, D. G., Alderman, F. R., and Blythe, W. B. (1980) Ala-Gly and Val-Asp-[Arg^8]-vasopressin: Bovine storage forms of arginine vasopressin with natriuretic activity. *Science* **207,** 893–896.
7. Valkna, A., Laidmä, E., Karelson, E., Zilmer, M., Juréus, A., and Langel, Ü. (1995) Effects of chimeric galanin receptor ligands on basal adenylate cyclase activity in rat ventral hippocampal membranes. *Protein Pept. Lett.* **2,** 267–274.
8. Östenson, C.-G., Zaitsev, S., Berggren, P.-G., Efendic, S., Langel, Ü., and Bartfai, T. (1997) Galparan: a powerful insulin releasing chimeric peptide acting at a novel site. *Endocrinology* **138,** 3308–3313.
9. Howl, J., Langel, Ü., Hawtin, S. R., et al. (1997) Chimeric strategies for the rational design of bioactive analogs of small peptide hormones. *FASEB J.* **11,** 582–590.
10. Howl, J., Yarwood, N. J., Stock, D., and Wheatley, M. (1996) Probing the V_{1a} vasopressin receptor binding site with pyroglutamate-substituted linear peptide antagonists. *Neuropeptides* **30,** 73–79.

11. Hällbrink, M., Saar, K., Östenson, C.-G., et al. (1999) Effects of vasopressin-mastoparan chimeric peptides on insulin release and GTPase activity. *Regul. Pept*. **82,** 45–51.
12. Longland, C. L., Mezna, M., Langel, Ü., et al. (1998) Biochemical mechanisms of calcium mobilisation induced by mastoparan and chimeric hormone-mastoparan constructs. *Cell Calcium* **24,** 27–34.
13. Farquhar, M., Soomets, U., Bates, R. L., Martin, A., Langel, Ü., and Howl, J. (2002) Novel mastoparan analogs induce differential secretion from mast cells. *Chem. & Biol*. **9,** 63–70.
14. Poyner, D. R., Soomets, U., Howitt, S. G., and Langel, Ü. (1998) Structural determinants for binding to CGRP receptors expressed by human SK-N-MC and Col 29 cells: studies with chimeric and other peptides. *Br. J. Pharmacol*. **124,** 1659–1666.
15. Wei, C. M., Kim C. H., Miller, V. M., and Burnett, J. C. (1993) Vasonatrin peptide —a unique synthetic natriuretic peptide and vasorelaxing peptide. *J. Clin. Invest*. **92,** 2048–2052.
16. Foran, S. E., Carr, D. B., Lipkowski, A. W., et al. (2000) A substance P-opioid chimeric peptide as a unique nontolerance-froming analgesic. *Proc. Natl. Acad. Sci. USA* **97,** 7621–7626.
17. Miscika, A., Lipkowski, A. W., Horvath, R., Porreca, F., Yamamura, H. I., and Hruby, V. J. (1994) Delta-opioid receptor-selective ligands—DPLPE-deltorphin chimeric peptide analogs. *Int. J. Pept. Protein Res*. **44,** 80–84.
18. Gupta, S., Pasha, S., Gupta, Y. K., and Bhardwaj, D. K. (1999) Chimeric peptide of Met-enkaphilin and FMRFa induces antinociception and attenuates development of tolerance to morphine antinociception. *Peptides* **20,** 471–478.
19. Rajott, D. and Ruoslahti, E. (1999) Membrane dipeptidase is the receptor for a lung-targeting peptide iodentified by in vivo phage display. *J. Biol. Chem*. **274,** 11593–11598.
20. Porkka, K., Laakkonen, P., Hoffman, J. A., Bernasconi, M., and Ruoslahti, E. (2002) A fragment of the HMGN2 protein homes to the nuclei of tumor cells and tumor endothelial cells in vivo. *Proc. Natl. Acad. Sci. USA* **99,** 7444–7449.
21. Laakkonen, P., Porkka, K., Hoffman, J. A., and Ruoslahti, E. A. (2002) Tumor-homing peptide with a targeting specificity related to lymphatic vessels. *Nat. Med.* **8,** 751 – 755.
22. Arap, W., Haedicke, W., Bernasconi, M., et al. (2002) Targeting the prostate for destruction through a vascular address. *Proc. Natl. Acad. Sci USA* **99,** 1527–1531.
23. Åkerman, M. E., Chan, W. C., Laakonen, P., Bhatia, S., and Ruoslahti, E. (2002) Nanocrystal targeting in vivo. *Proc. Natl. Acad. Sci. USA* **99,** 12617–12621.
24. Schally, A. V. and Nagy, A. (1999) Cancer chemotherapy based on targeting of cytotoxic peptide conjugates to their receptor on tumours. *Eur. J. Endocrinol*. **141,** 1–14.
25. Cheronis, J. C., Whalley, E. T., Allen, L. G., et al. (1994) Design, synthesis and in vitro activity of bis(succinimido)hexane peptide heterodimers with combined B_1 and B_2 antagonist activity. *J. Med. Chem*. **37,** 348–355.

26. Chan, D., Gera, L., Stewart, J., et al. (2002) Bradykinin antagonist dimer, CU210, inhibits the growth of human lung cancer cell lines by a "biased agonist" mechanism. *Proc. Natl. Acad. Sci. USA* **99,** 4608–4613.
27. Howl, J. and Wheatley, M. (1998) Biochemical pharmacology of total retro-inverso analogues of bradykinin and angiotensin II: Molecular recognition by G-protein-coupled receptors and angiotensin converting enzyme. *Lett. Peptide Sci*. **5,** 37–41.
28. Howl, J. and Wheatley, M. (1993) V_{1a} vasopressin receptors: Selective biotinylated probes. *Meth. Neurosci*. **13,** 281–296.
29. Howl, J., Yarwood, N. J., Davies, A. R. L., and Wheatley, M. (1996) Renal bradykinin and vasopressin receptors: Ligand selectivity and classification. *Kidney Int*. **50,** 586–592.
30. Halazy, S. (1999) G-protein coupled receptors bivalent ligands and drug design. *Expert Opin. Ther. Pat.* **9,** 431–446.
31. Bernatowicz, M. S., Matsueda, R., and Matsueda, G. R. (1986) Preparation of Boc-[S-(3-Nitro-2-pyridinesulfenyl)]-cysteine and its use for unsymmetrical disulfide bond formation. *Int. J. Pept. Protein Res*. **28,** 107–112.
32. Tam, J. P. and Spetzler, J. C. (1995) Chemoselective approaches to the preparation of peptide dendrimers and branched artificial proteins using unprotected peptides as building blocks. *Biomedical Peptides, Proteins & Nucleic Acids* **1,** 123–132.
33. Pooga, M., Juréus, A., Rezaei, K., et al. (1998) Novel galanin receptor ligands. *J. Peptide Res*. **51,** 65–74.
34. Isersky, C., Metzger, H., and Buell, D. N. (1975) Cell cycle-associated changes in receptors for IgE during growth and differentiation of a rat basophilic leukemia cell line. *J. Exp. Med.* **141,** 1147–1162.
35. Slaninová, J., Machová, A., Kuncarová, P., Maletínská, L., and Howl, J. (2001) Receptor affinity and biological activity of chimeric peptides combining bradykinin with arginine vasopressin of fibronectin-related 'RGD' sequences, in *2nd Hellenic Forum on Bioactive Peptides* (Cordopatis, P. A. ed.), Typorama, Patras, pp. 359–365.

II

Synthetic Methodologies and Applications

3

Modification of Peptides and Other Drugs Using Lipoamino Acids and Sugars

Joanne T. Blanchfield and Istvan Toth

Summary

The scientific literature is full of new small molecules and larger peptides identified as potential pharmaceutical agents for a variety of diseases. The majority of these compounds, however, will never progress into the clinic because of poor oral absorption and low metabolic stability. The development of practical, economic, and widely applicable systems to improve the bioavailability of drugs is a highly sought-after goal. The conjugation of a drug with lipid and/or sugar units represents one of the most important strategies being investigated in this new field of drug delivery. This chapter describes one method of introducing lipidic groups to drugs via lipoamino acids and also provides useful procedures for the efficient incorporation of sugar units into drugs, particularly peptide drugs, via solid phase synthesis.

Key Words: Drug delivery; lipoamino acids; carbohydrates; liposaccharides; bioavailability; peptide drugs; solid phase synthesis.

1. Introduction

The discovery of novel biologically active compounds is undoubtedly an important aspect of the process of drug development. However, unless these compounds can be administered efficiently, survive the many degradative mechanisms of the human body, and be absorbed across cell membranes and biological barriers to be delivered to their sites of action, they are like bullets without a gun. The field of drug delivery research has now been accepted as a vital and integral facet of modern drug discovery. The difficulties associated with drug delivery are particularly pronounced for peptide-based drugs and gene delivery therapies.

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

The large size and hydrophilic nature of these molecules mean that they are very unlikely to diffuse easily across cell membranes or biological barriers such as the epithelial layer of the GI tract or the blood–brain barrier (BBB). In addition, our bodies possess well-developed systems specifically designed to recognize and rapidly destroy these biomolecules. In both cases, the cleavage of a single bond in these complex molecules can render the drug inactive. Despite suffering from poor bioavailability caused by low absorption and low biological stability, peptides and oligonucleotides continue to be of increasing importance as pharmaceutical agents. In 2001, protein and peptide products (excluding vaccines) accounted for more than 10% of the total ethical pharmaceutical market. This market share is on the increase as the number of products in Phase II and III clinical trials is more than twice the number of products already on the market ***(1)***. Thus, the need for flexible, cost-effective drug delivery systems is great.

1.1. Lipids in Drug Delivery

Increasing the lipophilicity of many compounds, particularly peptides, by the addition of long aliphatic chains, or "lipid" moieties, has become a well-established method for improving bioavailability. This technique increases passive diffusion across epithelial barriers or can increase uptake into the lymphatic system ***(2–4)***. Passive diffusion is by far the major route for compounds crossing epithelial barriers and so is the most important target for improving bioavailability. The presence of lipid groups has also been shown to have a protective effect for peptides, significantly increasing their biological half-lives ***(5)***.

The technique for the introduction of lipidic groups into compounds that is favored by the authors is the use of lipoamino acids (Laas) (**Fig. 1**). These are α-amino acids with alkyl sidechains that can be varied in length, substitution, and degree of unsaturation simply by altering the bromoalkane used in the synthesis ***(6)***. The benefits of Laas include the ability to couple them using standard peptide coupling procedures, making them ideal for use in solid phase syntheses. Their bifunctional nature means they can be introduced into a peptide at any point in the sequence and the number of Laas introduced is easily varied.

1.2. Sugars in Drug Delivery

The addition of sugar moieties to peptides and other drug candidates with poor bioavailability has three major benefits:

- The polyhydroxylated nature of sugars provides an efficient and biocompatible way of altering the physicochemical properties of a drug, particularly increasing water solubility.

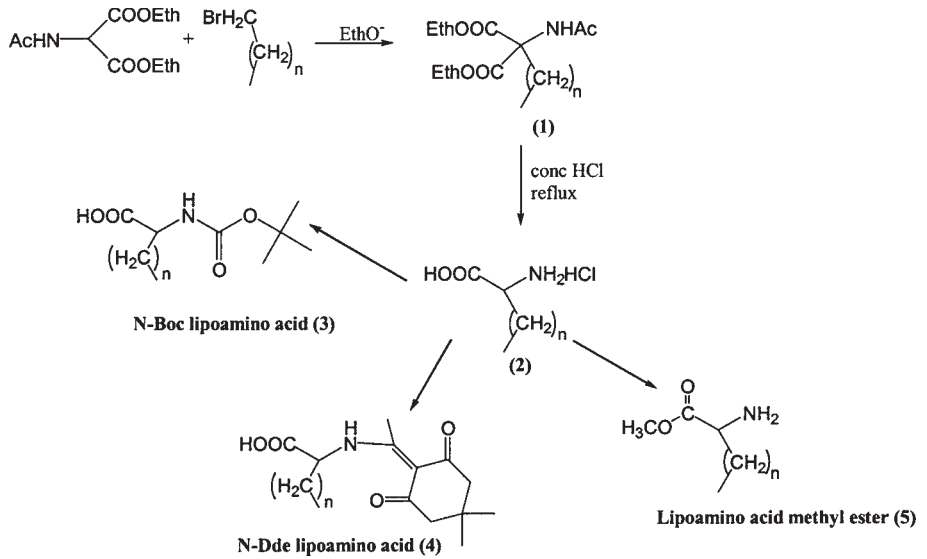


Fig. 1. The synthesis of protected lipoamino acids.

- Sugar units provide the opportunity of utilizing specific, active, or facilitated transport pathways across some biological barriers, particularly the BBB *(7)*.
- The recognition of some carbohydrate structures allows for the targeting of drugs to particular cell lines or organs.

Synthetic chemistry involving sugars has long had a reputation for being difficult. Establishment of techniques that allow the incorporation of sugar moieties into solid phase peptide synthesis is therefore an important breakthrough. The authors have developed such techniques *(8)* and, some examples of these are detailed below. Sugar units can be conjugated to the N-terminus of peptides or to any amine or alcohol group in a molecule via a spacer group such as succinic acid (**Fig. 2**). To incorporate a glucose unit into the C-terminus of a peptide, the sugar unit needs to be attached to the solid phase resin initially and the first amino acid coupled to a carboxyl group on the sugar (**Fig. 3**).

1.3. Lipids and Sugars in Drug Delivery

The ability of the polyhydroxylated carbohydrate moieties to offset the solubility problems caused by lipid conjugation has prompted researchers to combine the two strategies in drug delivery system development. As a consequence of all the factors previously discussed, lipid and/or sugar conjugation represents an important aspect of drug delivery research. A comprehensive review of these strategies at work in modern drug development can be found in our

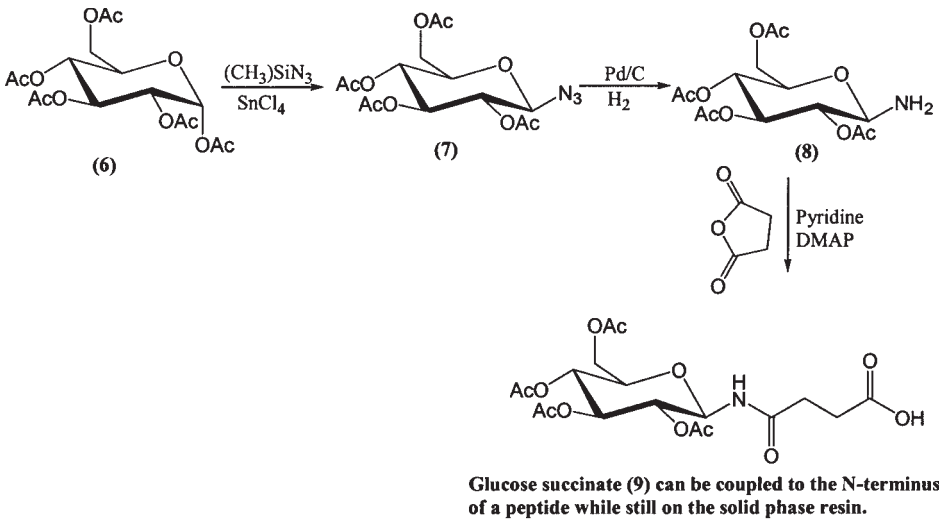


Fig. 2. Synthesis of glucopyranosyl succinamic acid.

two *Current Medicinal Chemistry* papers ***(9,10)***. These papers discuss direct chemical modifications to drug molecules to improve their bioavailability. However, liposaccharides such as 21 have been shown to enhance the absorption of the aminoglycoside gentamicin simply by coadministration ***(11)***. In this case **21** acts as an absorption enhancer (**Fig. 4**).

1.4. Dendrimer Formation Using Lipids and Sugars

The design of dendrimer scaffold molecules is becoming increasingly important in the area of gene and vaccine delivery. The lipoamino acids described earlier can be used to generate such dendrimer scaffolds as compound 22 **(Fig. 5)** named by the authors as the lipophilic core peptide (LCP) delivery system. This LCP system can be generated on solid phase resin using standard peptide synthetic protocols. It has already proved extremely effective as a scaffold for a new form of vaccine design. In this case, multiple copies of different antigenic peptide sequences can be conjugated at the positions designated as R. Our experiments suggest that these peptides exhibit excellent immunogenicity even without a chemical adjuvant ***(12,13)***. The LCP system can also be used to generate a polycationic dendrimer useful for the delivery of DNA into cells. For this application the R-positions are taken by further lysine groups bearing two primary amines that can be protonated, giving rise to the polycation. When this dendrimer is exposed to short oligonucleotide sequences they form complexes that can be successfully delivered into the cells of the eye ***(14)***.

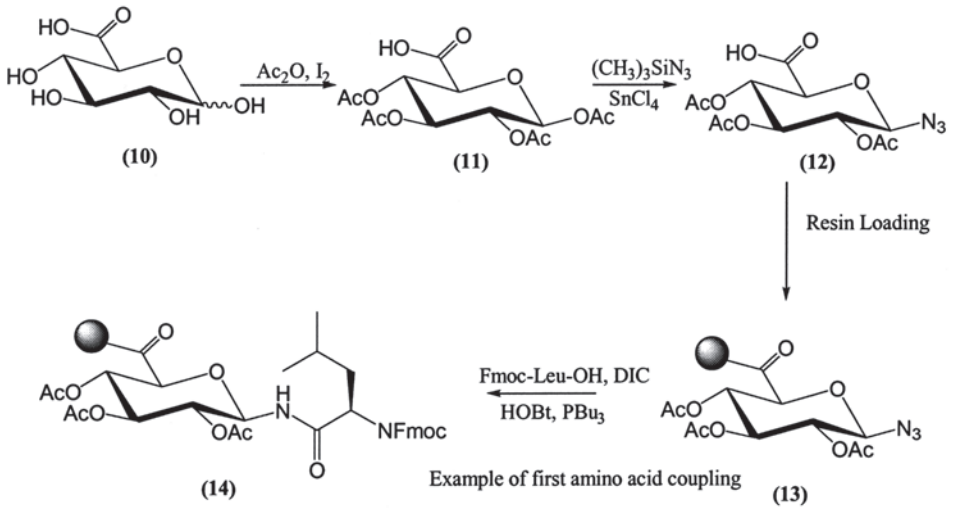


Fig. 3. Synthesis of glycopeptides on solid phase resin.

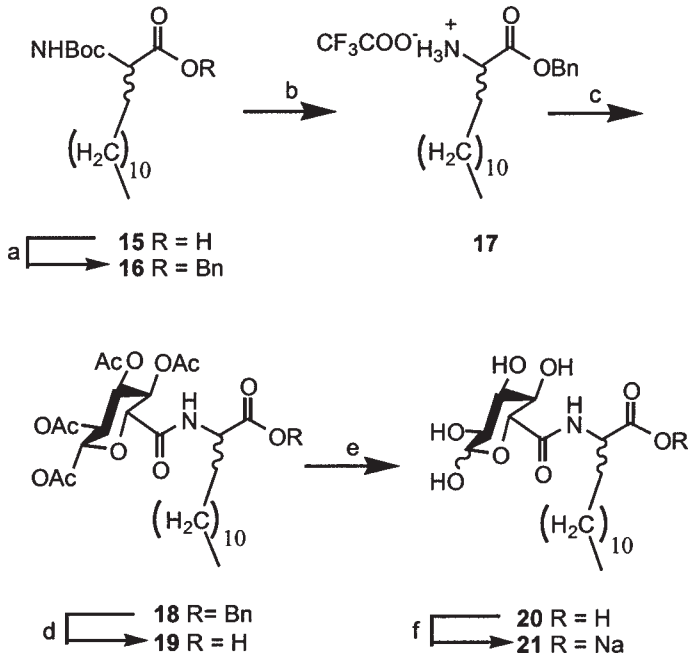
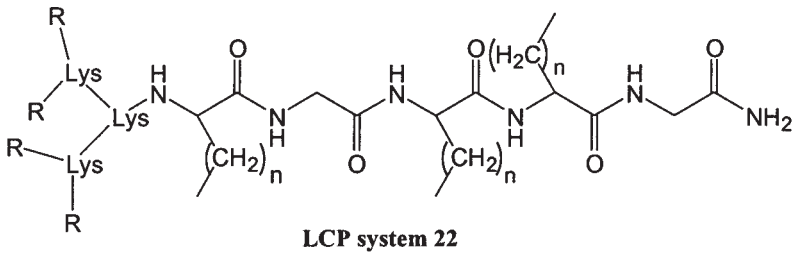
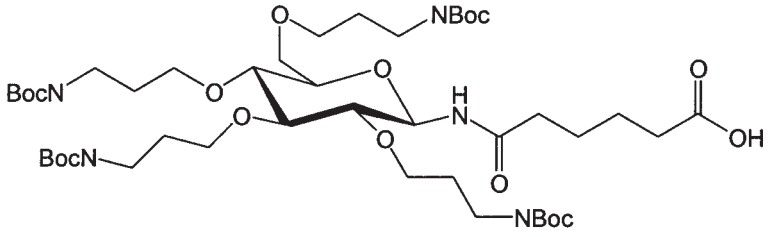


Fig. 4. Synthesis of absorption enhancer **21**. Reagents and conditions: (a) (1) Cs_2CO_3, MeOH; (2) BnCl, DMF, 70°C, 4 h; (b) CF_3COOH, DCM, 1 h; (c) **11**, HBTU, DIEA, THF, o/n; (d) H_2, 10% Pd/C, THF, o/n; (e) (1) 0.1 *M* NaOMe, MeOH, 2 h; (2) Amberlite IR-120(H^+); (f) $NaHCO_3$.



LCP system 22



Protected sugar based dendrimer 23

Fig. 5. Lipids and sugars as dendrimer scaffolds.

Carbohydrate structures can also act as a scaffold for dendrimer formation. The polysubstituted sugar 23 **(Fig. 5)** has been synthesized and shown that it can act as a scaffold onto which peptide sequences can be assembled to form another novel vaccine system ***(15)***. The development of this carbohydrate scaffold technology and its application to vaccine and gene delivery is under way. The synthesis of 23 is complex; a complete account of the procedures can be found in **ref.** ***15***.

1.5. Conclusion

Lipids and sugars, and the combination of both, are popular and successful strategies in drug delivery. This chapter provides details of just a handful of ways these moieties can be incorporated into molecules. There are, of course, many others and the choice of lipidic or carbohydrate group depends very heavily on the functional groups available for manipulation in the drug molecule and the size of that molecule. Drug delivery is always a delicate balance between adequate lipophilicity to cross bilipid layers and water solubility to aid in formulation.

2. Materials

2.1. Solvents

1. Ethanol.
2. Dimethyl formamide (DMF).

3. Acetonitrile.
4. Ethyl acetate.
5. Dichloromethane (DCM).
6. Toluene.
7. Hexane.
8. Diethyl ether.
9. Methanol.
10. Ethylene glycol dimethyl ether (DME).
11. Tetrahydrofuran (THF).
12. 2-Methyl-2-propanol (t-butanol).

2.2. Acids and Bases

1. 4-Dimethyl aminopyridine (DMAP).
2. Triethylamine.
3. NH_3.
4. NaOH.
5. Sodium metal.
6. Cs_2CO_3.
7. LiOH.
8. Concentrated HCl.
9. 1 *M* HCl.
10. Acetic acid.
11. Citric acid.

2.3. Aqueous Solutions

1. Saturated NaCl.
2. Saturated $NaHCO_3$.
3. 1 *M* $KHSO_4$.
4. 1 *M* $Na_2S_2O_3$.

2.4. Other Reagents

1. Diethyl acetamido malonate.
2. 1-Bromo alkanes.
3. Nitrogen or argon gas.
4. Hydrogen gas.
5. Di-*tert*-butyldicarbonate.
6. Anhydrous $MgSO_4$.
7. Dimedone.
8. Acetic anhydride.
9. 2% Hydrazine hydrate in methanol.
10. 12.1% Hydrazine hydrate in methanol.
11. Thionyl chloride.
12. 1,2,3,4,6-penta-*O*-acetyl-α-D-glucopyranose.

13. Trimethylsilyl azide.
14. Tin (IV) chloride.
15. 10% Pd/C.
16. Celite.
17. Succinic anhydride.
18. D-Glucuronic acid.
19. I_2.
20. 4-Methylbenzhydrylamine hydrochloride, polymer-bound (MBHA) resin.
21. 1-Hydroxybenzotriazole hydrate (HOBt).
22. 2-Dimethylaminoisopropyl chloride hydrochloride (DIC).
23. Tri-*n*-butyl phosphine.
24. Benzyl chloride.
25. Ion exchange resin (Amberlite IR-120 (H^+, Aldrich Chemical Company, Milwaukee, WI).

3. Methods

3.1. Lipids in Drug Delivery Methods

3.1.1. α-Lipoamino Acids **(2)**

Sodium metal (166 mmol) is dissolved in 100 mL ethanol under an inert atmosphere in a round-bottom flask fitted with a water condenser. Diethyl acetamido malonate (138 mmol) is then added, followed by 193 mmol the 1-bromoalkane of choice (e.g., 1-bromodecane will provide C_{12} Laa). The solution is refluxed overnight under an inert atmosphere. Upon cooling, the mixture is poured onto 600 mL crushed ice and stirred. The aminodiester product precipitates and is collected on a Buchner funnel and air-dried. The crude product is then refluxed overnight in a 200-mL solution of HCl:DMF (9:1). Upon cooling, the product precipitates and is collected on a Buchner funnel, washed with ice water, and air-dried to afford the α-amino acid hydrochloride typically in >90% crude yield.

Experimenters' notes: The refluxing acid step hydrolyzes the ester and the amide groups and decarboxylates the diacid **(Fig. 1)**. When long alkyl chains are employed (e.g., >14 carbons), this may require longer than 24 h. The reaction should be monitored by electrospray mass spectrometry (ESMS) to ensure it is complete. Again, for long alkyl chain Laas it is imperative that a very large flask be used in the acid reflux step, and this reaction should be monitored carefully. If insufficient headspace over the reaction is allowed it can explode out of the condenser.

3.1.2. N-Boc Protection **(3)**

Laa hydrochloride synthesized above (96.3 mmol) is suspended in a 500-mL solution of 2-methyl-2-propanol:water (2:3) and the pH adjusted to 13.0 with

5 *M* sodium hydroxide. Di-*t*-butyldicarbonate (144 mmol) in 50 mL 2-methyl-2-propanol is added. The solution is stirred overnight, maintaining the pH at 13.0. The mixture is then diluted with 200 mL water and solid citric acid added to pH 3.0. The mixture is extracted with ethyl acetate (5 × 150 mL) and the combined extracts dried ($MgSO_4$) and evaporated to yield the crude product. The *N*-Boc Laa is recrystallized from boiling acetonitrile.

Experimenters' notes: 2-Methyl-2-propanol = *tert*-butanol. The pH of this reaction should be checked regularly and adjusted back to 13.0, as it will change considerably.

3.1.3. Use of Laas in Fmoc-Based Synthesis

Fmoc-protected Laas are extremely insoluable in most solvents, including DMF, so they are not very useful in solid phase synthesis. However, using the 1-(1-hydroxyethylidene)-5,5-dimethylcyclohexane-1,3-dione (Dde) protecting group on the amine group provides an excellent alternative for use in Fmoc-based syntheses (**Fig. 4**). **Subheadings 3.1.4.–3.1.6.** constitute a description of the synthesis of the protecting group from readily available dimedone, a procedure for attaching the protecting group to the amino acid, and the deprotection protocol compatible with solid phase Fmoc synthesis.

3.1.4. 2-(1-Hydroxyethylidene)-5,5-Dimethylcyclohexane-1,3-Dione (Dde-OH)

Dimedone (15 g, 0.107 mol) is dissolved in 100 mL dichloromethane. DMAP (2.61 g, 0.021 mol) and 30 mL triethylamine (0.22 mol) are added and the mixture stirred for 10 min. Acetic anhydride (12.2 mL, 0.129 mol) is added and the mixture stirred for 2 d under an inert atmosphere. The solvent is removed *in vacuo* by co-evaporation with toluene and the crude product is taken up in 300 mL ethyl acetate and washed with 5% HCl solution (3 × 300 mL). The organic solution is then dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to produce an oil that is filtered through a column of silica using hexane/ethyl acetate 3:2 solution as an eluent to afford the title compound as pale yellow crystals (~60% yield).

3.1.5. 2-(4,4-Dimethyl-2,6-Dioxocyclohex-1-ylidene) Ethylaminododecanoic Acid (Dde-C_{12}-OH **[4]***)*

2-Amino-D,L-dodecanoic acid hydrochloride (C_{12}-Laa) (12.3 g, 48.7 mmol) and 9.76 g Dde-OH (53.6 mmol) are suspended in 200 mL ethanol. Triethylamine (12.3 g, 122 mmol) is added and the mixture refluxed under an inert atmosphere for 2 d. The solvent is removed *in vacuo* and the crude product taken up in 250 mL ethyl acetate and washed with 5% HCl (3 × 200 mL) then dried ($MgSO_4$). The solvent is removed to afford a solid that is triturated with diethyl ether.

Repeated trituration with diethyl ether affords pure product as a white solid (~55% yield).

3.1.6. Deprotection of N-Dde-Laas

The N-Dde protected Laas can be coupled to a growing peptide on Rink amide resin using standard coupling techniques. When the desired coupling efficiency is reached the protecting group is removed by treatment with 2% hydrazine hydrate in DMF (three washes of 5 min) followed by efficient rinsing of the resin with DMF as usual.

3.1.7. Methyl Ester of Laas **(5)**

Thionyl chloride (5 mL, 69 mmol) is added dropwise to 50 mL methanol cooled to 0°C. After 10 min, the HCl salt of a Laa (21 mmol) is added and dissolved with stirring. The ice bath is removed when dissolved and the reaction stirred overnight. Solvent is removed by evaporation to afford the crude solid (>90% yield).

3.1.8. Hydrolysis of Methyl Ester

If hydrolysis of the methyl ester of the Laas is required, this can be achieved by dissolving the compound in DME and water and adding LiOH (~5 equivalents) and allowing the reaction to stir at room temperature overnight. The reaction mixture is then acidified to pH 4.0 with 5% citric acid solution and extracted into ethyl acetate (3 × 20 mL). The organic extracts are washed with saturated NaCl, dried with $MgSO_4$, filtered, and solvent removed. The purification of the resulting free acid depends on the compound to which the Laa has been coupled.

3.2. Sugar Drug Delivery Methods

3.3. Solution Conjugation

3.3.1. 2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosyl Azide **(7)**

1,2,3,4,6-Penta-*O*-acetyl-α-D-glucopyranose (5.00 g, 12.8 mmol) is dissolved in 10–15 mL dry CH_2Cl_2 under an inert atmosphere. Trimethylsilyl azide (4.24 mL, 32.1 mmol) and 0.75 mL tin (IV) chloride (6.41 mmol) are added and the reaction allowed to proceed for 18 h at room temperature. The solution is diluted with CH_2Cl_2 and washed twice with saturated $NaHCO_3$ and once with saturated NaCl. The organic phase is dried, filtered, and concentrated to leave 3.92 g of the title compound as a white solid (82%). FAB MS ($C_{14}H_{19}N_3O_9$) 373.11 m/z (%): 331 $[M\text{-}N_3]^+$ (13), 396 $[M+Na]^+$ (7), 506 $[M+Cs]^+$ (100). ^{1}H-NMR ($CDCl_3$): δ 1.99, 2.02, 2.07, 2.09 (4s, 12H, 4OAc), 3.81 (m, 1H, H-5),

4.15, 4.27 (2m, 2H, H-6 and H-6'), 4.65 (d, 1H, H-1, $J_{1,2}$ = 8.8 Hz), 4.94 (t, 1H, H-2), 5.09, 5.21 (2t, 2H, H-3 and H-4).

3.3.2. 2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosylamine **(8)**

The glycosyl azide ***(7)*** (7.80 g, 20.9 mmol) is dissolved in 50 mL dry methanol and the solution placed under an inert atmosphere. 10% Pd on activated charcoal (300 mg) is added and the reaction maintained under a constant pressure of H_2 with vigorous stirring for 24 h at room temperature. The suspension is filtered through celite and the solvent removed under reduced pressure to yield 6.70 g of the title compound as a colorless syrup (92%). The residue can be used immediately for the preparation of glycosyl amide ***(9)***. FAB MS ($C_{14}H_{21}NO_9$) 347.12 m/z (%): 169 (79), 331 $[M\text{-}NH_2]^+$ (34), 370 $[M+Na]^+$ (100).

3.3.3. N-(2,3,4,6-Tetra-O-Acetyl-β-D-Glucopyranosyl) Succinamic Acid **(9)**

The glycosylamine product of the previous reaction (6.70 g, 19.3 mmol) is dissolved in 80 mL dry CH_2Cl_2 at 0°C. Dry pyridine (7.81 mL, 96.5 mmol) and 90 mg DMAP (0.74 mmol) are added. A solution of 4.83 g succinic anhydride (48.3 mmol) in dry CH_2Cl_2 is then added dropwise over 30 minutes and the reaction allowed to proceed for 30 min at 0°C, and a further 3 h at room temperature. The solution is then washed three times with 1 *M* HCl, once with saturated NaCl, then dried, filtered, and concentrated to leave a white foam. The title compound is crystallized from ethyl acetate-hexane (6.61 g, 77%). FAB MS ($C_{18}H_{25}NO_{12}$) 447.14 m/z (%): 331 $[M\text{-}NHCO(CH_2)_2COOH]^+$ (93), 448 $[M+H]^+$ (68), 470 $[M+Na]^+$ (100), 580 $[M+Cs]^+$ (55). ^{1}H-NMR ($CDCl_3$): δ 2.03, 2.05, 2.06, 2.08 (4s, 12H, 4OAc), 2.49 (t, 2H, CH_2CON), 2.70 (m, 2H, CH_2COO), 3.80 (m, 1H, H-5), 4.05 (dd, 1H, H-6), 4.28 (dd, 1H, H-6'), 4.92 (dd, 1H, H-2), 5.05 (dd, 1H, H-4), 5.24 (dd, 1H, H-1, $J_{1,2}$ = 9.0 Hz), 5.28 (dd, 1H, H-3), 6.52 (d, 1H, NH), 8.01 (br s, 1H, COOH).

3.4. Sugars on Solid Phase Resin

3.4.1. 1,2,3,4-Tetra-O-Acetyl-β-D-Glucopyranuronic Acid **(11)**

D-Glucuronic acid (6.00 g, 30.9 mmol) is suspended in 85 mL acetic anhydride and stirred at 0°C. Iodine (425 mg, 1.67 mmol) is added slowly and stirring continued for 2 h at 0°C, then a further 1 h at room temperature. The solution is then cooled to 0°C, 30 mL dry methanol added dropwise, and the solution allowed to stand for 18 h at room temperature. The reaction is concentrated and the residue taken up in 100 mL CH_2Cl_2, extracted, washed with 1 *M* $Na_2S_2O_3$, dried, filtered, and concentrated. The white residue is taken up in a mixture of ether, hexane, and $CHCl_3$ and concentrated again. Addition of ether, filtering, and

washing yield 9.20 g of the title compound as a fine white powder (82%). FAB MS ($C_{14}H_{18}O_{11}$) 362.08 m/z (%): 303 $[M\text{-}OAc]^+$ (34), 325 $[M\text{-}OAc+Na]^+$ (64), 385 $[M+Na]^+$ (86), 407 $[M+2Na\text{-}H]^+$ (100). ^{1}H-NMR ($CDCl_3$): δ 2.02, 2.03, 2.04, 2.11 (4s, 12H, 4OAc), 4.24 (m, 1H, H-5), 5.13 (m, 1H, H-2), 5.29 (m, 2H, H-3 and H-4), 5.79 (d, 1H, H-1, $J_{1,2}$ = 7.3 Hz). Anal. Calc. for $C_{14}H_{18}O_{11}$: C, 46.41; H, 4.97. Found: C, 46.24; H, 5.01.

3.4.2. 2,3,4-Tri-O-Acetyl-1-Azido-1-Deoxy-β-D-Glucopyranuronic Acid **(12)**

The protected glucopyranuronic acid from the previous reaction (2.00 g, 5.52 mmol) is dissolved in 40 mL dry CH_2Cl_2 under an inert atmosphere. Trimethylsilyl azide (1.87 mL, 13.8 mmol) and 0.32 mL tin (IV) chloride (2.76 mmol) are then added to the solution. The reaction is allowed to proceed for 18 h at room temperature under an inert atmosphere. The solution is then diluted with CH_2Cl_2, washed twice with 1 *M* $KHSO_4$, dried, filtered, and concentrated to yield 1.64 g of the title compound as a solid white foam (86%). FAB MS ($C_{12}H_{15}N_3O_9$) 345.08 m/z (%): 303 $[M\text{-}N_3]^+$ (35), 368 $[M+Na]^+$ (100), 390 $[M+2Na\text{-}H]^+$ (22). ^{1}H-NMR ($CDCl_3$): δ 2.02, 2.04, 2.07 (3s, 9H, 3OAc), 4.17 (d, 1H, H-5), 4.75 (d, 1H, H-1, $J_{1,2}$ = 8.8 Hz), 4.96 (m, 1H, H-2), 5.28 (m, 2H, H-3 and H-4). Anal. Calc. for $C_{12}H_{15}N_3O_9$: C, 41.74; H, 4.35; N, 12.17. Found: C, 41.65; H, 4.40; N, 12.19.

3.4.3. Immobilization of 2,3,4-Tri-O-Acetyl-1-Azido-1-Deoxy-β-D-Glucopyranuronic Acid onto MBHA Resin **(13)**

MBHA resin (1.96 g, 0.59 mmol/g substitution, 1.16 mmol) is swelled in dry DMF and deprotected as usual. The azide product of the previous reaction (1.21 g, 3.51 mmol) is dissolved in a minimum volume of dry DMF and cooled to 0°C. HOBt (0.54 g, 3.51 mmol) is added followed by 0.55 mL DIC (3.51 mmol) and the solution stirred for 10 min at 0°C. The solution and 0.61 mL *N*,*N*-diisopropylethylamine (DIEA) (3.51 mmol) are then added to the resin and the suspension mixed for 3 h. The resin is drained, washed, and dried *in vacuo* over KOH to constant weight (2.33 g, 99%, 0.50 mmol/g loading).

Experimenter's notes: The same procedure can be used to attach the sugar to Rink amide resin for Fmoc-based synthesis. The Rink amide resin must be deprotected as usual before coupling of the sugar.

3.4.4. First Amino Acid Attachment **(14)**

The dried resin from the previous reaction (100 mg, 0.63 mmol/g loading, 63 μmol) is transferred to a solid-phase reaction vessel, washed, and swelled in dry DMF under an inert atmosphere for at least 1 h. The resin is then washed several times with dry THF and left as a slurry. The first protected amino acid (0.25 mmol) is dissolved in dry THF, HOBt is added, and the solution cooled

to 0°C. DIC (40 μL, 0.25 mmol) is then added and the entire solution added to the resin, followed by 47 μL tri-*n*-butylphosphine (0.19 mmol). The suspension is mixed for 18 h under an inert atmosphere. The resin is then drained and washed well with DMF.

Experimenter's notes: The coupling efficiency can be monitored by the standard ninhydrin test. If recoupling is necessary it can be performed as a standard amino acid coupling, as the excess tri-*n*-butylphosphine in the above reaction reduces the azide of the sugar to the amine.

3.4.5. Removal of O-Acetyl Protecting Groups

Solid phase: De-*O*-acetylation is carried out after removal of the terminal *N*-protecting group but prior to peptide cleavage. The drained resin-peptide is suspended in 2 mL of 12.5% (v/v) hydrazine hydrate in methanol and the suspension mixed for 18 h at room temperature. The resin is then drained and washed well with DMF before preparing for cleavage as usual.

Solution phase: De-*O*-acetylation is achieved by suspending the lyophilized peptide or compound in 5 mL of approx 5% (w/w) NH_3 in methanol. The reaction is allowed to proceed for 2 h at room temperature. The solvent is then removed under reduced pressure and the compound taken up in 95% (v/v) acetic acid and lyophilized.

3.5. Synthesis of Liposaccharide Absorption Enhancer 21

3.5.1. 1,2,3,4-Tetra-O-Acetyl-β-D-Glucopyranuronic Acid

Method A: D-glucuronic acid (6.00 g, 31 mmol) is suspended in 85 mL acetic anhydride and stirred at 0°C. Iodine (427 mg) was added slowly over half an hour and stirring continued at 0°C for 2 h, then for 1 h at 25°C. The solution is then cooled to 0°C, 30 mL dry methanol added dropwise, and then the solution allowed to stir for 18 h at 25°C. The solution is concentrated and taken up in 100 mL DCM, washed with 1 *M* sodium thiosulphate solution (2 × 100 mL), dried ($MgSO_4$), filtered, and evaporated *in vacuo* to yield 8.3 g of a pale, brown gum. The crude product is taken up in a mixture of 120 mL ether, hexane, and chloroform (1:1:1) and crystallized upon concentration of the mixture. The crystalline product is triturated with ether to give 6.17 g of the title compound (55%) ***(4)*** as a white powder.

Method B: Acetic 1,2,3,4-tetra-β-acetyl-β-D-glucopyranuronic anhydride (2.04 g, 5.0 mmol) is dissolved in a mixture of 40 mL THF and 10 mL water and stirred overnight. The solvent is removed *in vacuo* to afford acetylated glucuronic acid as a white powder in quantitative yield: TLC R_f 0.41 (MeOH/DCM 1:3, ceric ammonium sulfate dip); mp 149°C; ESI-MS, *m/z*: 385 $[M+Na]^+$, 380 $[M+NH_4]^+$, 303 $[M\text{-}OAc]^+$; ^1H NMR (500 MHz, $CDCl_3$) δ 8.57 (1H, br s, COOH),

5.77 (1H, d, $J_{1,2}$ 7.6, H-1), 5.28 (2H, m, H-3 and H-4 overlapping), 5.10 (1H, t, J 8.0, H-2), 4.23 (1H, d, $J_{4,5}$ 9.0, H-5), 2.08, 2.01, 2.01, 2.00 (12H, 4s, 4Ac); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.1, 169.8, 169.7, 169.3, 169.0, 91.2, 72.3, 71.8, 70.1, 68.6, 20.6, 20.4, 20.4, 20.4; FT-IR 2954, 1728, 1427, 1373, 1211, 1042 cm^{-1}.

*3.5.2. 2-(*Tert-*Butoxycarbonylamino)-D,L-Tetradecanoic Acid Benzyl Ester* **(16)**

2-(*tert*-Butoxycarbonylamino)-D,L-tetradecanoic acid (*see* **Subheading 3.1.2.** for synthesis) (4.10 g, 11.9 mmol) is dissolved in 300 mL methanol. The solution is neutralized with a 20% solution of Cs_2CO_3 (approx 30 mL) and the solvent removed *in vacuo*. 50 mL DMF is added and the solvent removed *in vacuo*. Addition and evaporation of 50 mL DMF is repeated to produce the solid cesium salt, which was dried under reduced pressure. The cesium salt is dissolved in 150 mL DMF and 1.66 g benzyl chloride (1.51 mL, 13.1 mmol) added. The mixture is stirred at 70°C for 4 h. The solvent is removed *in vacuo* and the crude product taken up in 200 mL ethyl acetate and washed with acid (5% HCl, 2 × 100 mL), saturated bicarbonate (2 × 100 mL), and 100 mL NaCl. The organic layer is dried ($MgSO_4$), filtered, and evaporated to yield 3.88 g of the title compound (75%) ***(16)*** as an oil: ESI-MS, *m/z*: 566 $[M+Cs]^+$, 451 $[M+NH_4]^+$, 434 $[M+H]^+$; ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.34 (5H, m, Ar-H), 5.16 (3H, m, Ar-CH_2 overlapping with CONH), 4.34 (1H, m, α-CH), 1.9-1.5 (2H, m, β-CH_2), 1.44 (9H, s, $C(CH_3)_3$), 1.26 (20H, m, $10CH_2$), 0.89 (3H, t, J 10.9, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.7, 155.3, 135.4, 128.4, 128.2, 128.1, 79.6, 66.7, 53.4, 32.5, 31.8, 29.5, 29.4, 29.3, 29.1, 28.2, 25.1, 22.6, 14.0; HRMS calc. for $[M+H]^+$ 434.3264, found 434.3270.

3.5.3. 1-D,L-(Benzyloxycarbonyl)Tridecyl-Ammonium Trifluoroacetate **(17)**

Benzyl ester ***(16)*** (471mg, 1.09 mmol) is dissolved in 20 mL DCM/TFA (1:1) and stirred for 1 h. The solvent is removed *in vacuo* and the residue dissolved in acetonitrile/water (1:1) and lyophilized to produce 463 mg of the title compound (95%): ESI-MS, *m/z*: 334 $[M+H]^+$; ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.34 (5H, m, Ar-H), 5.24 (1H, d, J 12.1, Ar-CH), 5.13 (1H, d, J 12.1, Ar-CH), 4.01 (1H, m, α-CH), 1.91 (2H, m, β-CH_2), 1.23 (20H, m, $10CH_2$), 0.90 (3H, t, J 6.4, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.7, 134.5, 128.7, 128.6, 128.5, 68.1, 53.1, 31.9, 30.5, 29.6, 29.6, 29.5, 29.3, 29.1, 29.0, 24.6, 22.7, 14.1; HRMS calc. for $[M+H]^+$ 334.2740, found 334.2731.

3.5.4. N-[1-D,L-(Benzyloxycarbonyl)Tridecyl]-1,2,3,4-Tetra-O-Acetyl-β-D-Glucopyranuronamide **(18)**

1,2,3,4-Tetra-*O*-acetyl-β-D-glucopyranuronic acid (2.19 g, 6.0 mmol) is dissolved in 50 mL THF. DIEA (1.6 g, 2.1 mL, 12.1 mmol) and HBTU (2.50 g,

6.6 mmol) are added and the mixture stirred for 10 min at room temperature. A solution of 1,2,3,4-tetra-*O*-acetyl-β-D-glucopyranuronic acid (2.46 g, 5.5 mmol) in 50 mL THF ***(17)*** is then added and the mixture stirred overnight. The solvent is removed *in vacuo* to yield an oil that is taken up in 200 mL DCM and washed with 5% HCl (2 × 100 mL), saturated bicarbonate (2 × 100 mL), 50 mL water, and 50 mL NaCl. After drying ($MgSO_4$) the solvent is removed *in vacuo* to give an oil that is purified by silica flash column chromatography (ethyl acetate/hexane, 1:2) to yield 2.81 g of the title compound (75%) ***(18)***: m.p. 65–67°C; TLC R_f 0.37 (ethyl acetate/hexane 1:2, ninhydrin dip); ESI-MS, *m/z*: 700 $[M+Na]^+$, 695 $[M+NH_4]^+$, 678 $[M+H]^+$, 618 $[M-AcOH+H]^+$; ^1H NMR (300 MHz, $CDCl_3$) δ 7.32 (5H, m, Ar-H); 6.82 (1H, m, amide NH); 5.76 (1H, m, H-1); 5.33-5.07 (5H, m, overlapping Ar-CH_2, H-2, H-3, H-4); 4.55 (1H, m, α-CH); 4.09 (1H, m, H-5); 2.10, 2.02, 1.99 (12H, 3s, 4 Ac overlapping); 1.9-1.5 (2H, m, β-CH_2); 1.23 (20H, m, 10CH_2); 0.85 (3H, t, J 6.6, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.7, 171.6, 169.7, 169.7, 169.3, 169.1, 169.1, 168.7, 168.5, 165.6, 165.4, 135.2, 135.1, 128.5, 128.3, 128.2, 128.1, 91.2, 91.2, 73.0, 72.8, 71.9, 71.9, 70.1, 68.9, 68.7, 67.0, 67.0, 51.9, 51.8, 32.2, 32.1, 31.8, 29.5, 29.3, 29.2, 29.0, 29.0, 24.9, 24.7, 22.6, 20.6, 20.5, 20.5, 20.4, 14.0; HRMS calc. for $[M+H]^+$ 678.3484, found 678.3506.

3.5.5. N-[1-D,L-(Carboxy)Tridecyl]-1,2,3,4-Tetra-O-Acetyl-β-D-Glucopyranuronamide **(19)**

Benzyl ester (2.68 g, 4.0 mmol) ***(18)*** is dissolved in 50 mL THF. Pd/C catalyst (0.5 g, 10% Pd on C) is added and the solution stirred overnight under an atmosphere of H_2. The mixture is filtered through celite and the solvent evaporated *in vacuo* to yield 2.31 g of the title compound (99%) ***(19)*** as an oil: TLC R_f 0.16 (ethyl acetate/hexane 4:1, ninhydrin dip); ESI-MS, *m/z*: 588 $[M+H]^+$, 528 $[M-AcOH+H]^+$; ^1H NMR (300 MHz, $CDCl_3$) δ 8.72 (1H, br s, COOH), 6.83 (1H, m, amide NH); 5.79 (1H, d, J 7.6, H-1); 5.32-5.11 (3H, m, overlapping H-2, H-3, H-4); 4.52 (1H, m, α-CH); 4.16 (1H, m, H-5); 2.13, 2.04, 2.00 (12H, 3s, 4 Ac overlapping); 1.9-1.5 (2H, m, β-CH_2); 1.24 (20H, m, 10CH_2); 0.86 (3H, t, J 6.2, CH_3); ^{13}C NMR (75 MHz, $CDCl_3$) δ 175.6, 175.4, 169.9, 169.6, 169.5, 169.3, 169.2, 168.8, 168.7, 166.1, 91.3, 91.2, 72.9, 72.8, 72.0, 71.9, 70.2, 69.0, 68.8, 51.8, 51.8, 32.0, 31.8, 29.6, 29.3, 29.2, 25.1, 25.0, 22.6, 20.7, 20.5, 14.0; HRMS calc. for $[M+H]^+$ 588.3014, found 588.3026.

3.5.6. N-[1-D,L-(Carboxy)Tridecyl]-α,β-D-Glucopyranuronamide **(20)**

Compound (2.20 g, 3.7 mmol) ***(19)*** is dissolved in 50 mL methanol. Sodium methoxide solution (50 mL, 0.2 *M*) is added and the solution stirred at room temperature for 2 h. The mixture is neutralized with ion-exchange resin (Amberlite IR-120 $[H^+]$), filtered, and then the solvent removed *in vacuo*. The product is lyophilized from acetonitrile/water (1:1) to yield the title compound (1.49 g,

95%) ***(20)***: ESI-MS, *m/z*: 442 $[M+Na]^+$, 420 $[M+H]^+$; ^1H NMR (300MHz, d_6-DMSO) δ 8.1-7.5 (1H, m, amide NH); 4.4-2.9 (m); 1.9-1.5 (2H, m, β-CH_2); 1.23 (20H, m, $10CH_2$); 0.85 (3H, t, J 6.3, CH_3); HRMS calc. for $[M+H]^+$ 420.2591, found 420.2592.

3.5.7. Sodium N-[1-D,L-(Carboxy)Tridecyl]-α,β-D-Glucopyranuronamide **(21)**

The free acid (0.701 g, 1.67 mmol) ***(20)*** is suspended in 50 mL water and 3.34 mL sodium bicarbonate (0.5 *M* solution, 1.67 mmol) added. The mixture is sonicated until a clear solution is obtained. The reaction mixture is lyophilized to give 0.734 g of the title compound ***(21)*** in quantitative yield. ESI-MS, *m/z*: 442 $[M+Na]^+$, 420 $[M+H]^+$; ^1H NMR (500 MHz, d_6-DMSO) δ 7.6 (1H, m, amide NH); 4.4-2.9 (m); 1.68 (1H, m, β-CH); 1.51 (1H, m, β-CH); 1.23 (20H, m, $10CH_2$); 0.85 (3H, t, J 6.8, CH_3); HRMS calc. for $[M+H]^+$ 420.2591, found 420.2621; HPLC-MS, CN t_R = 5.82 min, ODS t_R = 5.25, 5.69 min (overlapping), 95%.

References

1. Loffet, A. (2002) Peptides as drugs: is there a market? *J. Peptide Sci.* **8,** 1–7.
2. Garzon-Aburbeh, A., Poupaert, J. H., Claesen, M., Dumont, P., and Atassi, G. (1983) 1,3-Dipalmitoylglycerol ester of chlorambucil as a lymphotropic, orally administrable antineoplastic agent. *J. Med. Chem.* **26,** 1200–1203.
3. Garzon-Aburbeh, A., Poupaert, J. H., Claesen, M., and Dumont, P. (1986) A lymphotropic prodrug of L-dopa: Synthesis, pharmacological properties, and pharmacokinetic behavior of 1,3-dihexadecanoyl-2-[(S)-2-amino-3-(3,4-dihydroxyphenyl)propanoyl]propane-1,2,3-triol. *J. Med. Chem.* **29,** 687–691.
4. Porter, C. J. (1997) Drug delivery to the lymphatic system. *Crit. Rev. Ther. Drug Carrier Sys.* **14,** 333–393.
5. Toth, I., Flinn, N., Hillery, A., Gibbons, W. A., and Artursson, P. (1994) Lipidic conjugates of luteinizing hormone releasing hormone (LHRH) and thyrotropin releasing hormone (TRH) that release and protect the native hormones in homogenates of human intestinal epithelial (Caco-2) cells. *Int. J. Pharm.* **105,** 241–247.
6. Gibbons, W. A., Hughes, R. A., Charalambous, M., et al. (1990) Synthesis, resolution and structural elucidation of lipidic amino acids and their homo- and hetero-oligomers. *Liebigs Ann. Chem.* **12,** 1175–1183.
7. Polt, R. and Palian, M. M. (2001) Glycopeptide analgesics. *Drugs of the Future* **26,** 561–576.
8. Malkinson, J. P., Falconer, R. A., and Toth, I. (2000) Synthesis of C-terminal glycopeptides from resin-bound glycosyl azides via a modified Staudinger reaction. *J. Org. Chem.* **65,** 5249–5252.
9. Wong, A. and Toth, I. (2001) Lipid, sugar and liposaccharide based delivery systems. *Curr. Med. Chem.* **8,** 1123–1136.

10. Blanchfield, J. T. and Toth, I. (2004) Lipids, sugars and liposaccharides in drug delivery 2: an update. *Curr. Med. Chem.* **11,** 2375–2382.
11. Ross, B. P., DeCruz, S. E., Lynch, T. B., Davis-Goff, K., and Toth, I. (2003) Design, Synthesis and evaluation of a liposaccharide drug delivery agent: application to the gastrointestinal absorption of gentamicin. *J. Med. Chem.* **47,** 1251–1258.
12. Horvath, A., Olive, C., Wong, A., et al. (2002) Lipoamino acid-based adjuvant carrier system: Endhanced immunogenicity of group A streptococcal peptide epitopes. *J. Med. Chem.* **45,** 1387–1390.
13. Olive, C., Batzloff, M., Horvath, A., et al. (2003) Potential of lipid core peptide technology as a novel self-adjuvanting vaccine delivery system for multiple different synthetic peptide immunogens. *Infection and Immunity* **71,** 2373–2383.
14. Wimmer, N., Marano, R. J., Kearns, P. S., Rakaczy, E. P., and Toth, I. (2002) Syntheses of polycationic dendrimers on lipophilic peptide core for complexation and transport of oligonucleotides. *Biorg. Med. Chem. Lett.* **12,** 2635–2637.
15. McGeary, R. P., Jablonkai, I., and Toth, I. (2001) Carbohydrate-based templates for synthetic vaccines and drug delivery. *Tetrahedron* **57,** 8733–8742.

4

Synthesis of Linear, Branched, and Cyclic Peptide Chimera

Gábor Mezö and Ferenc Hudecz

Summary

Chimeric peptides are unnatural constructs consisting of bioactive compounds from at least two different peptide(s) and/or protein(s) or two sequences from different parts of the same protein. Such multifunctional peptide combinations are prepared to enhance the biological activity or selectivity of their components. New biological effects can also be achieved with the chimera. In this chapter the synthesis of three different types of chimeric peptides will be described. In a linear chimera, two peptide epitopes from different parts of glycoprotein D (gD) of herpes simplex virus (HSV) are combined. A branched chimera, built from linear peptides, consists of tuftsin oligomers with immunostimulatory activity and an epitope peptide of HSV gD. The third compound is a cyclic chimeric molecule, where α-conotoxin GI as a host peptide is modified by the incorporation of a core epitope from HSV gD as a guest sequence.

Key Words: Chimeric peptide; cyclopeptide; solid phase peptide synthesis; thioether bond; disulfide bridge; herpes simplex virus glycoprotein D; peptide epitope; oligotuftsin; carrier molecule.

1. Introduction

The phrase "chimeric peptide" is widely used in the literature for unnatural peptide constructs consisting of bioactive compounds from at least two different peptides and/or proteins or two sequences from different parts of the same protein. Such multifunctional peptide combinations are prepared to enhance the biological activity or selectivity of their components. New biological effects can also be achieved with chimeric constructs. A chimeric peptide composed of

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

the endogenous opioid tetrapeptide (endomorphin-2) and neuropeptide substance P (SP) sequences activates both μ-opioid and SP receptors, resulting in opioid-dependent analgesia in the rat spinal cord without the development of tolerance upon repeated administration ***(1)***. Incorporation of an Arg-Gly-Asp (RGD) sequence into various proteins recognized by many integrins causes them to acquire cell adhesion activity ***(2,3)***. Cellular uptake of different peptides can be increased by their attachment to cell-penetrating peptides such as penetratin ***(4)***, HIV-1 Tat ***(5)***, antennapedia ***(6)***, or transportan, which is itself a chimeric peptide combining the amino-terminal of galanin with mastoparan (*see* Chapter 5, ***[7]***). Preparation of synthetic antigens is one of the main areas for the application of peptide chimeras. Combination of sequential B-cell epitopes or T-cell and B-cell epitopes in a single molecule is frequently used for the analysis of epitope regions ***(8)*** or for the development of synthetic vaccines ***(9)***. In some cases one of the peptide sequences (guest) is utilized for stabilization of the structure of the host peptide to increase its bioactivity ***(10,11)***. According to a new approach, natural cyclic compounds (e.g., conotoxins) with well-characterized 3D structures are used as a host scaffold peptide/protein to accept a guest sequence. The guest sequence, derived from another peptide or protein, partially replaces the primary structure of the host to provide a different biological activity. This strategy takes advantage of the conformational similarity of the host and guest sequences ***(12,13)***.

There are many examples of the preparation of chimeric peptides using recombinant DNS technology. Such protocols utilize *E. coli* or other vectors for the expression of unnatural peptide combinations. This strategy is very well suited to the preparation of long linear peptides containing only amide bonds ***(5,10)***. However, there are some limitations to common recombinant strategies. In some cases the isolation of recombinantly expressed peptide chimeras was unsuccessful because of solubility problems ***(14)***. The recombinant DNS procedure is also not suitable when unnatural amino acids (e.g., D-amino acids ***[11]***, ε-amino caproic acid, or β-alanine ***[8]***) or phosphopeptides **(3)** are present in the target sequence. Similarly, the synthesis of branched peptides ***(15)*** or cyclopeptides ***(8,11–13)*** cannot be achieved by this technique. Hence, chemical synthesis must be used for the preparation of nonlinear peptides and chimeras containing unnatural building blocks. Interestingly, there are also many examples of the chemical synthesis of linear chimeric peptides, especially shorter sequences.

Linear peptides are the most common example of peptide chimeras. In this strategy two different peptide/protein sequences are tandemly linked. These compounds can be chemically assembled using stepwise ***(16,17)*** or convergent solid phase synthetic procedures ***(18).*** The length and/or the amino acid composition of desired sequences may dictate different approaches. The other possibility is the attachment of monomeric peptides by the formation of nonpeptide

bonds (e.g., disulfide bridge ***[19]*** or thioether bond ***[20]***) or by some other chemical ligation technique ***[21]***). These approaches, especially the latter ones, may allow the use of purified unprotected peptides, resulting in a better yield use (only) in many cases. Chimeric cyclopeptides are typically produced by formation of disulfide bridges from linear precursors ***(2,11–14)***.

2. Materials

1. *Resins for peptide synthesis:* 4-methylbenzhyrylamine-resin (MBHA) for Boc chemistry and 4-(2',4'-dimethoxyphenyl-aminomethyl)phenoxy (Rink)-amide resin for Fmoc chemistry (Novabiochem; Laufelfingen, Switzerland).
2. *Boc-amino acid derivatives:* Boc-Acp-OH, Boc-Ala-OH, Boc-Arg(Tos)-OH·EtOAc, Boc-Asn-OH, Boc-Asp(OcHex)-OH, Boc-Cys(Meb)-OH, Boc-Gly-OH, Boc-Glu (OcHex)-OH, Boc-Leu-OH·H_2O, Boc-Lys(Fmoc)-OH, Boc-Lys(ClZ)-OH, Boc-Nle-OH·DCHA, Boc-Phe-OH, Boc-Pro-OH, Boc-Ser(Bzl)-OH, and Boc-Val-OH (Reanal; Budapest, Hungary).
3. *Fmoc-amino acid derivatives:* Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(O^{t}Bu)-OH, Fmoc-Cys(Acm)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-Pro-OH, Fmoc-Tyr(^{t}Bu)-OH, and Fmoc-Val-OH (Novabiochem).
4. *Cleavage and neutralization reagents:* trifluoroacetic acid (TFA)[1], liquid hydrogen fluoride (HF)[1,2], *N,N*-diisopropylethylamine (DIEA),[1,3] and piperidine[2,3] (Fluka, Buchs, Switzerland or SDS, Peypin, France).
5. *Coupling reagents: N,N'*-dicyclohexylcarbodiimide (DCC)[2], *N,N'*-diisopropylcarbodiimide (DIC)[2], 1-hydroxybenzotriazol (HOBt),[3] and *O*-benzotriazolyl-*N,N, N',N'*-tetramethyluronium hexafluorophosphate (HBTU)[4] (Fluka, or Applied Biosystems, SARL; Paris, France).
6. *Scavengers:* anisole[4], *m*-cresol[2], *p*-cresol[2], *p*-thiocresol[1], thioanisole[4], 1,2-ethandithiol (EDT)[2], dithiotreitol (DTT), and phenol[2] (Fluka or Sigma-Aldrich; St. Quentin Fallavier, France).
7. *Reagents for oxidation:* 5,5'-dithio-*bis*-(2-nitrobenzoic acid) (DTNB, Ellman reagent)[4], silver triflate (AgOTf)[4], thallium trifluoroacetate (Tl(tfa)$_3$)[2], iodine,[5] and guanidine HCl[5] (Fluka or Sigma-Aldrich).
8. *Solvents:* acetonitrile (AcN)[2,3], acetic acid (AcOH)[1], dichloromethane (DCM)[5], diethyl ether[3], *N,N*-dimethylformamide (DMF)[2], dimethylsulfoxide (DMSO)[4], ethanol,[3] and 1-methyl-2-pyrrolidone (NMP)[4] (Reanal or Novabiochem).
9. Chloroacetic acid pentachlorophenyl ester (ClAcOPcp) was prepared in our laboratory from chloroacetic acid[2,6] and pentachlorophenol[2,6] with aid of DCC (Fluka).

3. Methods

The methods described outline the synthesis of (1) linear, (2) branched, and (3) cyclic chimeric peptides. The linear peptide consists of two different epitope

1: corrosive, 2: highly toxic, 3: flammable, 4: irritant, 5: toxic, 6: hazardous, according to Fluka catalog.

regions of glycoprotein D from the HSV. The second example demonstrates the preparation of a branched peptide chimera using a sequential oligopeptide derived from IgG (tuftsin) that possesses immunostimulatory and/or adjuvant effects. An epitope from HSV gD-1 was attached to the oligomer tuftsin derivative, also considered as a carrier for biomedical applications, resulting in a branched peptide architecture. In the third compound an α-conotoxin peptide host molecule was combined with an epitope from HSV gD-1 as a guest sequence.

3.1. Synthesis of Chimeric Peptide Containing Two Epitope Regions From Glycoprotein D of HSV-1

The N-terminal sequence, residues 1–23, is a major immunodeterminant region of glycoprotein D (gD) from herpes simplex virus type 1 (HSV-1) ***(22)***. The core epitope with antigenic properties in this region was assigned to residues 9–21 in which both T- and B-cell epitopes are present (LKMADPNRFRGKD) ***(23)***. Another B-cell epitope was found in sequence 276 to 284 (SALLEDPVG) ***(24)***. The chimeric peptide was designed to join these two regions by connecting them through a flexible spacer (ε-amino caproic acid; Acp) with a cysteine residue available for future conjugation to carrier molecules with formation of a disulfide or thioether bridge. Since the replacement of methionine by norleucine in position 11 in the 9 to 21 sequence did not change the antigenicity of the epitope ***(25)***, this modification was also applied to the chimeric construct (*see* **Note 1**). However, during the synthesis of the H-^{9}LKNleADPNRFRGKD^{21}AcpC^{276}SALL EDPVG^{284}-NH_2 chimera, aspartimide formation at Asp-Acp was observed. This was due to Acp lacking a side-chain, which contributed to the side reaction appearance of α- and β-aspartyl peptides. To prevent this side reaction the 9 to 21 peptide was elongated at the C-terminus with a leucine residue also present in the natural sequence of gD protein. The H-^{9}LKNleADPNRFRGKDL^{22}AcpC^{276}SALL EDPVG^{284}-NH_2 chimeric peptide was then synthesized by step-wise solid phase methodology.

3.1.1. Solid Phase Synthesis of a Chimeric Peptide With Boc Strategy

The peptide was manually synthesized on MBHA resin (1.1 mmol/g capacity). The protected amino acid derivatives used for the synthesis were as follows: Boc-Gly-OH, Boc-Val-OH, Boc-Pro-OH, Boc-Asp(OcHex)-OH, Boc-Glu (OcHex)-OH, Boc-Leu-OH·H_2O, Boc-Ala-OH, Boc-Ser(Bzl)-OH, Boc-Cys (Meb)-OH, Boc-Acp-OH, Boc-Lys(ClZ)-OH, Boc-Arg(Tos)-OH·EtOAc, Boc-Phe-OH, Boc-Asn-OH, and Boc-Nle-OH·DCHA. Three equivalents (with respect to resin capacity) of protected amino acid derivatives were used for the coupling in the presence of equivalent amount of coupling reagents (DCC and HOBt) dissolved in a DCM-DMF mixture (*see* **Note 2**). The protocol of the synthetic cycle was:

1. Wash resin with DCM (three times, 0.5 min).
2. Remove Boc group with 33% TFA in DCM (2+ 20 min).
3. Wash resin with DCM (five times, 0.5 min).
4. Neutralize with 10% DIEA in DCM (three to four times/min).
5. Wash resin with DCM (four times, 0.5 min).
6. Couple 3 eq (amino acid derivatives-DCC-HOBt) (60 min).
7. Wash resin with DMF (two times, 0.5 min).
8. Wash resin with DCM (two times, 0.5 min).
9. Monitor the coupling steps by ninhydrin assay ***(26)*** or Bromophenol blue assay ***(27)*** in the case of proline to assess the presence of free amines.

In the case of negative results a new cycle was started. The cycle was repeated from the neutralization step when colorimetric tests indicated ineffective coupling. After assembling the peptide chain on the resin the N-terminal Boc-group was removed prior to cleavage with HF.

3.1.2. Cleavage and Deprotection With HF

After removal of the last Boc group the peptide-resin was neutralized and washed with DCM (four times) and ethanol (two times, 0.5 min). Before HF cleavage the resin was dried in a desiccator overnight. The cleavage of side chain protecting groups and the detachment of the peptide from the resin by HF was performed in a Teflon apparatus (Peptide Institute, Osaka, Japan). For 1 g of peptide resin 10 mL of liquid HF was used in the presence of 1 g *p*-cresol as scavenger and 0.1 g DTT to prevent oxidation of the thiol group of cysteine. The cleavage was continued for 90 min at 0°C. Shorter cleavage times may not be sufficient for the complete removal of the Tos protecting group of arginine residues and Meb from cysteine. After evaporation of HF (captured in a CaO-filled tower) the peptide was precipitated with cooled dry ether. The mixture of precipitated peptide and resin was filtered and washed three to four times with cooled dry ether (30–40 mL) to remove scavengers. The unprotected peptide was dissolved in 10% acetic acid solution and lyophilized.

3.1.3. Purification and Characterization of the Chimeric Peptide

1. *HPLC:* Analytical RP-HPLC was performed using a Phenomenex Jupiter C_{18} (250 × 4.6 mm, 5 μm, 300 Å) reverse phase column and a Waters system composed of No. 600 pump, No. 600 controller, and No. 490 programmable multiwavelength detector (Nihon Waters, Tokyo, Japan). Eluent A was 0.1% TFA in water and eluent B was 0.1% TFA in acetonitrile-water (80:20, v/v). A linear gradient of eluents (0–90% B in 45 min) with a 1 mL/min flow rate was used as the mobile phase. Three different wavelengths were applied for the detection of peaks (λ = 214, 254, and 280 nm).

   Purification of chimeric peptides was carried out on Phenomenex Jupiter C_{18} (250 × 10 mm, 10 μm, 300 Å) semipreparative column using the same eluents with

5 mL/min flow rate. Gradient was 10 to 50% B in 45 min. Detection was carried out at λ = 214, 254, and 280 nm.

2. *Amino acid analysis:* The amino acid composition of peptides was determined using a Beckman Model 6300 analyzer (Beckman, Fullerton, CA, USA). Prior to analysis, samples were hydrolyzed in 6 *M* HCl in sealed and evacuated tubes at 110°C for 24 h.
3. *Mass spectrometry:* Positive ion electrospray ionization mass spectrometry (ESI-MS) analysis was performed on a PE API 2000 triple quadrupole mass spectrometer (Sciex, Toronto, Canada). Spray voltage was set to 4.8 kV, and 30 V orifice voltage was applied. Samples were dissolved in a methanol-water (1:1, v/v) mixture containing 0.1% acetic acid, and 5 µL of sample was injected with a flow rate of 100 µL/min. The instrument was used in a Q_1 scan mode in the range of m/z 400–1700, with a step size of 0.3 amu and a dwell time of 0.5 ms. Other chimeric peptides in this study were purified and characterized in the same or a very similar way.

3.2. Synthesis of Branched Chimera Composed of Oligotuftsin Derivative With HSV Peptide Epitope

In this chimeric peptide construct the aim was to combine the carrier function and immunstimulatory activity of tuftsin derivatives with an epitope derived from HSV gD to achieve an increased antibody response. Tuftsin is a well-known natural tetrapeptide (TKPR) that has a pronounced effect on the immune system ***(28,29)***. Polymerized tuftsin (polytuftsin) is also considered as a carrier molecule that increased antibody levels against attached epitopes in mice ***(30,31)***. New, sequential oligopeptides based on repeated tuftsin derivatives (H-[Thr-Lys-Pro-Lys-Gly]$_n$-NH_2, where n = 2,4,6,8) were developed in our laboratory to eliminate the drawbacks of tuftsin derivatives produced by polymerization. These new, nontoxic, nonimmunogenic compounds have immunostimulatory activity and a minor chemoattractant effect on monocytes ***(32)***. An oligotuftsin derivative was used in this study for the synthesis of a peptide chimera containing an HSV peptide epitope.

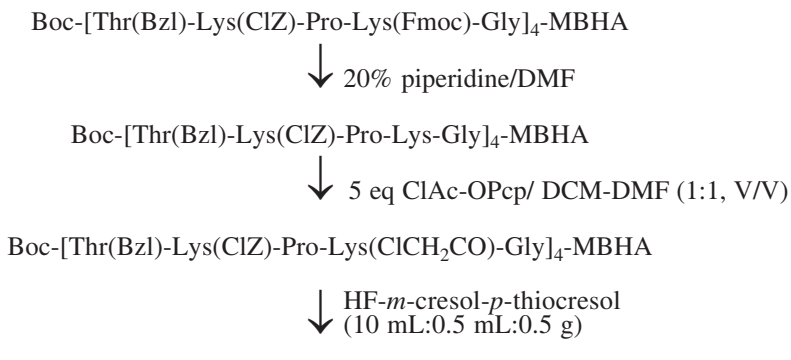
3.2.1. Synthesis of Chloroacetylated Oligotuftsin Derivatives H-[Thr-Lys-Pro-Lys(ClAc)-Gly]$_4$-NH_2 (see **Note 3**)

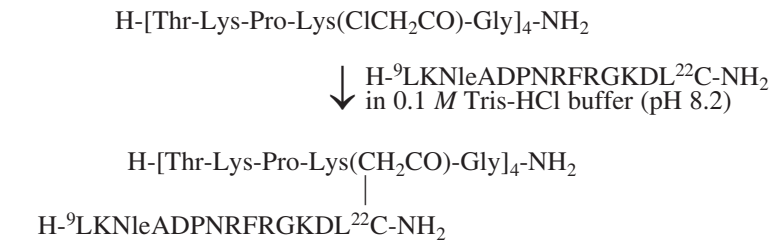
The oligotuftsin derivative containing four repeated units of the pentapeptide was prepared by solid phase synthesis on MBHA resin (1.04 mmol/g capacity) using Boc strategy. The following amino acid derivatives were used: Boc-Gly-OH, Boc-Lys(Fmoc)-OH, Boc-Pro-OH, Boc-Lys(ClZ)-OH, Boc-Thr(Bzl)-OH. The peptide was assembled using the protocol described in **Subheading 3.1.1.** The last Boc-Thr(Bzl)-OH was attached to the peptide-resin with a DIC/HOBt coupling method instead of DCC/HOBt. This strategy prevents the appearance of precipitated *N,N'*-dicyclohexylurea (DCU) that can be eliminated by TFA

treatment (Boc cleavage), but not when removing Fmoc groups with a piperidine/DMF mixture. After the attachment of last amino acid derivative the Fmoc groups were selectively cleaved from the ε-amino group of appropriate Lys residues with 20% piperidine/DMF mixture (2+2+5+10 min). The peptide-resin was washed eight times with DMF then two times with DCM (0.5 min each). Five equivalents of chloroacetic acid pentachlorophenyl ester (ClAc-OPcp) to each free amino group (20 eq total) dissolved in DCM-DMF (1:1, v/v) was used to incorporate the ClAc function (*see* **Note 4**). The N-terminal Boc group was removed prior to the final cleavage step. The cleavage of the peptide from the resin was carried out with liquid HF according to the protocol in **Subheading 3.1.2.** However, 0.5 g *p*-thiocresol and 0.5 mL *m*-cresol were used as scavengers as recommended in the literature for chloroacetylated peptide compounds ***(33)***. The crude peptide was purified by RP-HPLC and, after lyophilization, was characterized by analytical HPLC, mass spectrometry, and amino acid analysis. The pure product was obtained at 65% yield.

3.2.2. Coupling of the [Nle^{11}]-9-22-Cys Epitope Peptide From HSV gD-1 to Chloroacetylated Oligotuftsin Derivatives

Chloroacetylated oligotuftsin H-[Thr-Lys-Pro-Lys(ClAc)-Gly]$_4$-NH_2 was dissolved in 0.1 *M* Tris-HCl buffer (pH 8.2) in a peptide concentration of 1 mg/mL. The 20-mL tube used for the reaction was filled up fully and closed tightly. The epitope peptide synthesized by the procedure presented in **Subheading 3.1.1.** (H-^{9}LKNleADPNRFRGKDL^{22}C-NH_2) was added to the mixture in 1–2 mg aliquots every 15 min (*see* **Note 5**). In this way low concentrations of the epitope peptide prevented disulfide bond formation. The coupling was followed by analytical HPLC that showed the incorporation of the epitope peptide. The reaction was completed in 2 d using two equivalents of epitope peptide to each chloroacetyl group (*see* **Note 6**). The branched chimeric peptide was purified on RP-HPLC (60% yield) (*see* **Note 7**) and after lyophilization the purified product was characterized by analytical HPLC, amino acid analysis and mass spectrometry.





3.3. Synthesis of a Cyclic Chimera Consisting of α-Conotoxin GI as a Host and an HSV Epitope as a Guest Sequence

The 276-284 region of HSV gD-1 is highly hydrophilic and adopts a β-turn. The α-conotoxin GI also contains a β-turn in the 8–12 region, stabilized by two disulfide bridges at position 2–7 and 3–13 ***(34)***. Thus, the tetramer sequence of α-conotoxin, ^{8}Arg-His-Tyr-Ser^{12}, was replaced by ^{281}Asp-Pro-Val-Gly^{284} (DPVG), previously identified as the epitope core ***(35)***. Insertion of a tetrapeptide sequence representing a minimal B-cell epitope into the small structurally defined α-conotoxin scaffold resulted in a chimera that was immunogenic and able to induce DPVG-specific antibodies.

3.3.1. Synthesis of a Linear Precursor of a Cyclic Chimera Using the Fmoc Strategy

The synthesis of a linear precursor peptide of the HSV-α-[Tyr^{1}]-conotoxin chimera (*see* **Note 8**) was performed by solid phase methodology on an Applied Biosystems Mod. 431A peptide synthesizer (Applied Biosystems; Paris, France). Fmoc-based chemistry was used for the preparation of peptide using Rink-amide resin with 0.43 mmol/g loading. Four equivalents of Fmoc amino acid derivatives and HBTU/HOBt were used for coupling. The following amino acid derivatives were used in the synthesis: Fmoc-Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-Val-OH, Fmoc-Pro-OH, Fmoc-Cys(Acm)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Tyr(^{t}Bu)-OH. The synthetic protocol for Fmoc chemistry was as follows:

1. Wash resin with DMF (four times, 0.5 min).
2. Remove Fmoc with 20% piperidine in NMP (2 + 2 + 5 + 10 min).
3. Wash resin with DMF (eight times, 0.5 min).
4. Coupling (4 eq amino acid derivative–3.8 eq HBTU-4 eq HOBt) in NMP (30 min).
5. Wash resin with DMF (four times, 0.5 min).

After completing the sequence the N-terminal Fmoc group was cleaved with 20% piperidine in NMP. The peptide resin was washed with DMF (eight times, 0.5 min), DCM (three times, 0.5 min), and finally two times with ethanol (0.5 min). Prior to the cleavage of the peptide from the resin, it was dried in an evacuated desiccator for 3 h.

Peptide was cleaved from the resin with TFA mixtures containing different compositions of scavengers at room temperature: TFA-EDT (95:5, v/v), TFA-phenol-anisole-EDT (94:2:2:2, v/v) and Reagent K containing DTT instead of EDT (TFA-water-phenol-thioanisole-DTT = 33 mL:2 mL:2 g:2 mL:0.1 g) were compared in the cleavage reaction. By these procedures all protecting groups (^{t}Bu, Trt) were removed, except Acm. The quality of the crude product was not significantly influenced by the application of different cleavage mixtures. The crude peptides presenting a single peak in RP-HPLC were oxidized without any purification.

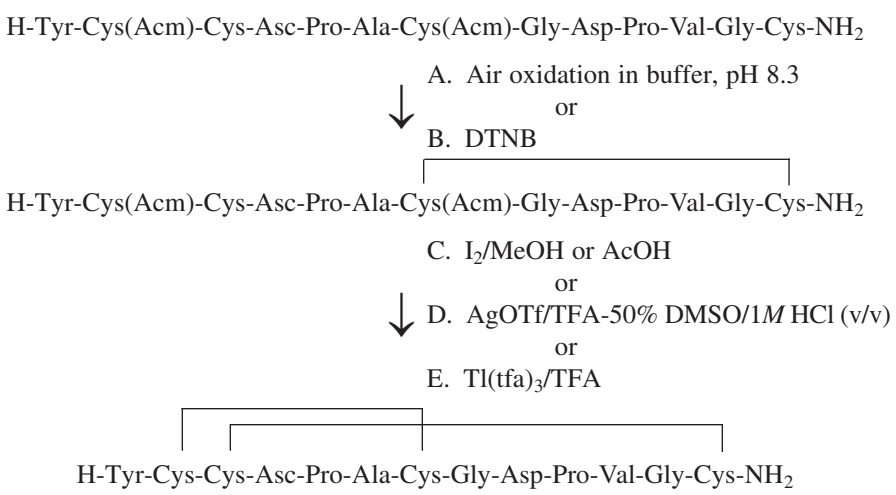
3.3.2. Cyclization (Disulfide Bridge Formation) of HSV-α-[Tyr^1]-Conotoxin Chimera

Five oxidation procedures were compared for the disulfide bridge formation.

1. **Air oxidation *(36)***, used for the first disulfide bridge formation of HSV-α-[Tyr^1]-conotoxin chimera, was carried out in 0.1 *M* potassium phosphate or ammonium acetate buffer at pH 8.3 in the absence or presence of guanidine HCl (1.0–5.0 *M*). Peptide concentration in solution was 0.1 mg/mL. The reaction was continued for 2 d at room temperature (RT) (*see* **Note 9**).
2. **Oxidation by Ellman reagent *(37)***: 1 eq (0.5 eq calculated for free thiol groups) 5,5'-dithio-*bis*(2-nitrobenzoic acid) (DTNB) was added to the peptide at 0.1 mg/mL in 20 m*M* ammonium bicarbonate (pH 8.7) at RT.
3. **Iodine oxidation *(38)***, was used for the second disulfide bond formation. Two equivalents of iodine were added to the peptide (0.1 mg/mL) in methanol or 95% acetic acid solution at RT. The iodine oxidation was continued for 2 d (*see* **Notes 10** and **11**).
4. **DMSO/1 M HCl oxidation *(39)***, was used for the second disulfide bond formation. The Acm groups were removed by treating 10 mg of the protected peptide in 2 mL TFA with 40 eq of AgOTf and in the presence of 30 µL anisole at 0°C for 1.5 h. The unprotected peptide was precipitated with ether, then collected by centrifugation. The oxidation was performed in 10% or 50% DMSO/1 *M* HCl solution (1 mg/mL peptide concentration) for 14 h at RT.
5. **Tl(tfa)$_3$/TFA oxidation *(40)***: The one loop containing Acm-protected peptide dissolved in TFA solution (25 mg/mL) containing 2% anisole was treated at 0°C for 1 h with 1.2 eq of Tl(tfa)$_3$. For preparative purposes the oxidation product was lyophilized and desalted on a Sephadex G10 column (1 × 20 cm) (Pharmacia, Uppsala, Sweden) equilibrated in 10% acetic acid. Pooled fractions were lyophilized and finally purified by RP-HPLC (*see* **Note 12**).

Fmoc-Tyr(^{t}Bu)-Cys(Acm)-Cys(Trt)-Asc(Trt)-
Pro-Ala-Cys(Acm)-Gly-Asp(O^{t}Bu)-Pro-Val-Gly-Cys(Trt)-R

↓ 1. 20% piperidine/DMF
or
2. 95% TFA-5% EDT



4. Notes

1. Replacement of methionine by norleucine is a common way to eliminate unwanted side reactions (oxidation and alkylation) that can be observed especially when using a Boc protection strategy. Owing to the similarity of Met and Nle side chains, the exchange usually has no significant influence on the structure of peptides.
2. In Boc chemistry the preferred solvent mixture for coupling is DCM-DMF (3:1 or 2:1, v/v). The resin swells enough for the effective interaction between reagents and resin bound peptide that solvates very well in DCM-DMF mixture.
3. To increase the enzymatic stability of oligotuftsin, N-terminal acetylation was applied in further experiments. Branching on Lys residues next to Gly residues increased the stability of the Lys-Gly bond in tryptic digestion.
4. Bromoacetyl groups cannot be used for thioether bond formation with cysteine residue in the presence of free amino groups. Only the chloroacetyl group is suitable for this purpose.
5. Inefficient coupling reaction occurred with the HSV-9-22 epitope peptide containing Cys at the N-terminus (H-Cys-^{9}LKNleADPNRFRGKDL^{22}-NH_2). This was caused by very rapid oxidation (dimer formation through disulfide bridge) that was not be prevented using a N_2 atmosphere in the presence of reducing agents. Rapid disulfide bridge formation can be explained by conformational or ionic interaction between two peptide chains. Before the coupling of epitopes with Cys at C- or N-terminus a study of oxidation is recommended. For effective conjugation, the half life of the free thiol groups should be more than 3 to 4 h in 0.1 *M* Tris-HCl buffer (pH 8.0–8.5) with a peptide concentration of 0.5 mg/mL.
6. The speed of coupling to oligotuftsin was dependent on the sequence of the epitope peptide and its capacity for oxidation resulting in disulfide bond containing homodimers. The incorporation of a spacer sequence dramatically increased the speed of the coupling to oliogotuftsin.

7. The efficacy of purification of the branched chimera containing oligotuftsin depended on the pattern of the HPLC chromatogram. A large difference between the retention time values of the epitope dimer and chimeric peptide favoured HPLC separation. However, when the peaks are close, reduction of the dimer with DTT could be helpful. The reduced epitope peptide can be separated by HPLC or by gel filtration on Sephadex G10. Sometimes a change of HPLC column from C_{18} to C_4 provided a good solution.
8. Glutamic acid was replaced by tyrosine in position one of α-conotoxin GI to prevent pyroglutamic acid formation at the N-terminal. This strategy also introduced an amino acid detectable at λ = 280 nm by UV.
9. Guanidine-HCl accelerates the rate of air oxidation, but more than 24 h were still needed for complete oxidation. DTNB oxidation was much faster, taking less than 10 min for the disulfide bond formation between ^{3}Cys and ^{13}Cys residues of HSV-α-[Tyr^{1}]-conotoxin chimera.
10. Peptides containing the Asp-Pro (DP) sequences are very sensitive to acidic conditions ***(41,42)***. To prevent cleavage of the Asp-Pro bond, careful handling is necessary even in dilute acids. Evaporation of acidic solvents in elevated temperature (over 20°C) or storage in acidic solution may lead to decomposition of Asp-Pro-containing peptides.
11. For the formation of the second disulfide between ^{2}Cys and ^{7}Cys, the method applying $Tl(tfa)_3$ obtained yields of over 90% and only a few percent of disulfide isomer were observed. Iodine oxidation in methanol gave no bicyclic product, while in 95% AcOH cleavage of the Asp-Pro bond was detected. In the case of DMSO/1 *M* HCl the oxidation took 14 h and side reactions occurred (Tyr deletion and appearance of disulfide isomers).
12. The Acm protecting group on Cys is not stable under electrospray mass spectrometry analysis. Besides the expected molecular masses (M) of M-71 and M-142, other peaks were also detected. When the source voltage was decreased from 340 V to 150 V the heterogeneity disappeared (Nermag R10-10ES apparatus coupled to an Analitica of Branford electrospray source).

Acknowledgments

This work was supported by grants from the French-Hungarian Intergovernmental Programme (Balaton) NP 867/94, from the Hungarian Spanish Intergovernmental Programme (5/1998), from the Hungarian National Science Fund (OTKA No. T014964, T032425, T037719), from the Ministry of Education (FKFP 0101/1997), and from the program entitled *Peptide Based Synthetic Antigens Against Infectious Diseases* COST Chemistry Action (D13/0007/00).

References

1. Foran, S. E., Carr, D. B., Lipkowski, A. W., et al. (2000) A substance P-opioid chimeric peptide as a unique nontolerance-forming analgesic. *Proc. Natl. Acad. Sci. USA* **97,** 7621–7626.

2. Vella, F., Hernandez, J.-F., Molla, A., Block, M. R., and Arlaud, G. J. (1999) Grafting an RGD motif onto an epidermal growth factor-like module: chemical synthesis and functional characterization of the chimeric molecule. *J. Pept. Res.* **54,** 415–426.
3. Gilbert, M., Shaw, W. J., Long, J. R., et al. (2000) Chimeric peptides of statherin and osteopontin that bind hydroxyapatite and mediate cell adhesion. *J. Biol. Chem*. **275,** 16213–16218.
4. Gil-Parrado, S., Assfalg-Machleidt, I., Fiorino, F., et al. (2003) Calpastatin exon 1B-derived peptide, a selective inhibitor of calpain: enhancing cell permeability by conjugation of with penetratin. *Biol. Chem*. **384,** 395–402.
5. Rusnati, M., Urbinati, C., Musulin, B., et al. (2001) Activation of endothelia cell mitogen activated protein kinase ERK(1/2) by extracellular HIV-1 Tat protein. *Endothelium* **8,** 65–74.
6. Ruzza, P., Donella-Deana, A., Calderan, A., et al. (2001) Antennapedia/HS1 chimeric phosphotyrosyl peptide: Conformational properties, binding capability to c-Fgr SH2 domain and cell permeability. *Biopolymers* **60,** 290–306.
7. Magzoub, M., Kilk, K., Eriksson, L, E., Langel, Ü., and Graslund, A. (2001) Interaction and structure induction of cell-penetrating peptides in the presence of phospholipid vesicles. *Biochim. Biophys. Acta* **1512,** 77–89.
8. Gomme, P. T., Thompson, P. E., Whisstock, J., Stanton, P. G., and Hearn, M. T. W. (1999) Characterization of epitope regions of thyrotropin b-subunit recognized by the monoclonal antibodies mAb279 and mAb299: a chimeric peptide approach. *J. Pept. Res.* **54,** 218–229.
9. Hel, Z., Johnson, L. M., Tryniszewska, E., et al. (2002) A novel chimeric Rev, tat, and Nef (Retanef) antigen as a component of an SIV/HIV vaccine. *Vaccine* **20,** 3171–3186.
10. Louis, J. M., Bewley, C. A., and Clore, G. M. (2001) Design and properties of N_{CCG}-gp41, a chimeric gp41 molecule with nanomolar HIV fusion inhibitory activity. *J. Biol. Chem.* **276,** 29,485–29,489.
11. Misicka, A., Lipkowski, A. W., Horvath, R., et al. (1994) Delta opioid receptor selective ligands; DPLPE-deltorphin chimeric peptide analoues. *Int. J. Pept. Prot. Res*. **44,** 80–84.
12. Mezö, G., Drakopoulou, E., Paál, V., Rajnavölgyi, É., Vita, C., and Hudecz, F. (2000) Synthesis and immunological studies of α-conotoxin chimera containing an immunodominant epitope from the 268–284 region of HSV gD protein. *J. Pept. Res.* **55,** 7–17.
13. Drakopoulou, E., Uray, K., Mezö, G., Price, M. R., Vita, C., and Hudecz, F. (2000) Synthesis and antibody recognition of mucin 1 (MUC-1)-α-conotoxin chimera. *J. Pept. Sci.* **6,** 175–185.
14. Hackeng, T. M., Rosing, J., Spronk, H. M. H., and Vermeer, C. (2001) Total chemical synthesis of human matrix Gla protein. *Protein Science* **10,** 864–870.
15. Mezö, G., de Oliveira, E., Krikorian, D., et al. (2003) Synthesis and comparision of antibody recognition of conjugates containing herpes simplex virus type 1 glycoprotein D epitope VII. *Bioconjug. Chem.* **11,** 1260–1269.

16. Stewart, J. M. and Young, J. D. (1984) *Solid Phase Peptide Synthesis* 2nd ed., Pierce Chemical Co., Rockford, IL.
17. Grant, G. A., ed. (1992) *Synthetic Peptides. A User's Guide*. W. H. Freeman and Company, New York.
18. Lloyd-Williams, P., Albericio, F., and Giralt, E. (1993) Convergent solid-phase peptide synthesis. *Tetrahedron* **49,** 11,065–11,133.
19. Andreu, D., Albericio, F., Solé, N. A., Munson, M. C., Ferrer, M., and Barany, G. (1994) Disulfide bond formation. *Methods Mol. Biol*. **35,** 91–169.
20. Robey, F. A. and Fields, R. L. (1989) Automated synthesis of bromoacetyl-modified peptides for preparation of synthetic peptide polymers, peptide-protein conjugates and cyclic peptides. *Anal. Biochem*. **177,** 373–377.
21. Dawson, P. E. and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation. *Ann. Rev. Biochem*. **69,** 925–962.
22. Welling-Wester, S., Scheffer, A.-J., and Welling, G. W. (1991) B and T cell epitopes of glycoprotein D of herpes simplex virus type 1. *Microbial. Immunol*. **76,** 59–68.
23. Welling-Wester, S., Feijbrief, M., Koedijk, D. G., et al. (1994) Anologues of peptide 9-21 of glycoprotein D of Herpes simplex virus and their binding to group VII monoclonal antibodies. *Arch. Virol.* **138,** 331–340.
24. Isola, V. J., Eisenberg, R. J., Siebert, G. H., Heilmanm, C. J., Wicox, W. C., and Cohen, G. H. (1989) Fine mapping of antigenic site II of herpes simplex virus glycoprotein D. *J. Virol.* **63,** 2325–2334.
25. Weijer, W. J., Drijfhout, J. W., Geerlings, H. J., et al. (1988) Antibodies against synthetic peptides of herpse simplex virus type 1 glycoprotein D and their capability to neutralize viral infectivity in vitro. *J. Virol.* **62,** 501–510.
26. Kaiser, E., Colescott, C. D., Bossinger, C. D., and Cook, P. I. (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34,** 595–598.
27. Krchnak, V., Vagner, J., Safar, P., and Lebl, M. (1988) Noninvasive continuous monitoring of solid-phase peptide synthesis by acid-base indicator. *Collect. Czech. Chem. Commun.* **53,** 2542–2548.
28. Fridkin, M. and Najjar, V. A. (1989) Tuftsin: its chemistry, biology, and clinical potential. *Crit. Rev. Biochem. Mol. Biol*. **24,** 1–40.
29. Siemion, I. Z. and Kluczyk, A. (1999) Tuftsin: On the 30-year anniversary of Victor Najjar's discovery. *Peptides* **20,** 645–674.
30. Trudelle, Y., Brack, A., Delmas, A., Pedoussaut, S., and Rivaille, P. (1987) Synthesis of a new carrier for immunization: polytuftsin. Two examples of its use with peptides selected in the hepatitis B surface antigen. *Int. J. Pept. Prot. Res*. **30,** 54–60.
31. Gokulan, K., Khare, S., and Rao, D. N. (1999) Increase in the immunogenicity of HIV peptide antigens by chemical linkage to polytuftsin ($TKPR_{40}$). *DNA Cell Biol.* **18,** 623–630.
32. Mezö, G., Kalászi, A., Reményi, J., et al. (2004) Synthesis conformation and immunoreactivity of new carrier molecules based on repeated tuftsin-like sequence. *Biopolymers* **73,** 645–656.

33. Ivanov, B. B. and Robey, F. A. (1996) Effective use of free thiols as scavangers for HF coctails to deprotect bromo- and chloroacetylated synthetic peptides. *Pept. Res.* **9,** 305–307.
34. Kobayashi, Y., Ohkubo, T., Kyogoku, Y., Sakakibara, S., Braun, W., and Go, N. (1989) Solution conformation of conotoxin GI determined by 1H nuclear magnetic resonans spectroscopy and distance geometry calculations. *Biochemistry* **28,** 4853–4860.
35. Hudecz, F., Hilbert, Á., Mezö, G., et al. (1993) Epitope mapping of the 273-284 region of HSV glycoprotein D by synthetic branched polypeptide-carrier conjugates. *Pept. Res.* **6,** 263–271.
36. Adamson, J. G. and Lajoie, G. A. (1994) Guanidine hydrochloride assists intramolecular disulfide bond formation in cysteine peptides. In *Peptides 1993* (Hodges, R. S. and Smith, J. A., eds.), ESCOM, Leiden, The Netherlands, pp. 44–45.
37. Ellman, G. L. (1959) Tissue sulfhydryl groups. *Arch. Biochem. Biophys*. **82,** 70–77.
38. Kamber, B. (1971) Cystine peptides from (S-acetamidomethyl-cysteine)-peptides through oxidation with iodine: synthesis of cyclo-L-cystine. *Helv. Chem. Acta* **54,** 927–930.
39. Tamamura, H., Otaka, A., Nakamura, J., et al. (1993) Disulfide bond formation in S-acetamidomethyl cysteine-containing peptides by the combination of silver trifluoromethansulfonate and dimethylsulfoxide/aqueous HCl. *Tetrahedron Lett*. **34,** 4931–4934.
40. Fujii, N., Otaka, A., Funakoshi, S., Bessho, K., and Yajima, H. (1987) Studies on peptides. CLI. Synthesis of cystine-peptides by oxidation of S-protected cysteine-peptides with thallium(III) trifluoroacetate. *Chem. Pharm. Bull*. **35,** 2339–2347.
41. Landon, M. (1977) Cleavage at aspartyl-prolyl bonds. *Methods Enzymol*. **47,** 145–149.
42. Skribanek, Z., Mezö, G., Mák, M., and Hudecz, F. (2002) Mass spectrometric and chemical stability of the Asp-Pro bond in herpes simplex virus epitope peptides compared with X-Pro bonds of related sequences. *J. Pept. Sci.* **8,** 398–406.

5

Synthesis of Cell-Penetrating Peptides for Cargo Delivery

Margus Pooga and Ülo Langel

Summary

Cell-penetrating peptides (CPPs) have served as vehicles for the delivery of different molecules and particles into cells. The efficiency and noninvasive nature of peptide-mediated cellular transduction provides a promising tool for biomedical research. Studies of cell-penetrating constructs that deliver oligonucleotides, peptides, and proteins have elucidated several important cellular signaling mechanisms. Here we briefly describe the major classes of cell-penetrating peptides. We also present various strategies used to couple different cargoes to CPPs.

Key Words: Cell-penetrating peptide; noninvasive cellular delivery; penetratin; Tat; transportan; pVEC; MAP; protein transduction; oligonucleotide transport; plasmid delivery.

1. Introduction

The first strategy to facilitate the cellular uptake of hydrophilic macromolecules with biological function, e.g., proteins and oligonucleotides, was their modification with cationic ***(1,2)*** or hydrophobic ***(3)*** moieties. The discovery and design of peptides that enter cells without using endocytosis or viral vectors yielded a more powerful tool for the cellular delivery of nonpermeable compounds ***(4)***. The term "cell-penetrating peptide" (CPP) was introduced for these sequences in order to distinguish their mechanism of cellular uptake, which was prevented neither by low temperatures nor by the inhibition of endocytosis. Initial investigations discovered CPPs that were truncated cationic fragments of naturally occurring proteins able to translocate from the medium into cells. Pioneering studies identified the penetratin sequence ***(5)*** from the Antennapedia

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

Table 1
Examples of Cell-Penetrating Peptides

Class	Peptide	Sequence[a] *(Ref.)*
Cationic	Penetratin (pAntp)	RQIKIWFQNRRMKWKK ***(5)***
	Tat (HIV-1)	GRKKRRQRRRPPQ ***(6)***
	Rev (HIV-1)	RRRRNRTRRNRRRVR ***(15)***
	Arg_9	RRRRRRRRR ***(14)***
	pIsl-1	RVIRVWFQNKRCKDKK ***(66)***
Hydrophobic	MTS	AAVALLPAVLLALLAP ***(10)***
	Integrin h-region	VTVLALGALAGVGVG ***(18)***
Amphipathic	MAP	KLALKLALKALKAALKLA ***(13)***
	HA-2	GLFGAIAGFIEGGWTGMIDG ***(19)***
Proline-rich/ antimicrobial	Bac (1-15)	RPIRPRPPRLPRPRP ***(23)***
	Bac (15-24)	PRPLPFPRPG ***(23)***
	Buforin 2	TRSSRAGLQFPVGRVHRLLRK ***(24)***
	Magainin 2	GIGKFLHSAKKFGKAFVGEIMNS ***(24)***
Bipartite/ chimeric	Transportan	GWTLNSAGYLLGKINLKALAALAKKIL ***(12)***
	SP-NLS	MGLGLHLLVLAAALQGAWSQPKKKRK ***(11)***
	Pep-1	KETWWETWWTEWSQPKKKRKV ***(26)***
	pVEC	LLIILRRRIRKQAHAHSK ***(27)***
	PrP	ANLGYWLLALFVTMWTDVGLCKKRPKP ***(28)***

[a]Peptides are amidated at the C-terminal.

homeodomain of *Drosophila* and the Tat ***(6)*** peptide from HIV Tat protein ***(7, 8)***. Studies that determined the efficiency of these CPPs for the cellular delivery of various cargoes soon followed ***(9)***. In parallel it was demonstrated that hydrophobic peptides corresponding to the h-region of membrane protein signal sequences were also cell-translocation competent ***(10)***. Soon the first designed CPPs appeared that comprised both cationic and hydrophobic regions ***(11,12)*** or were of an amphipathic nature, such as MAP peptide ***(13)***. Accordingly, a common feature of CPP-mediated cellular delivery is that the coupling of cargoes to penetrant peptides renders them more cationic and hydrophobic. The CPPs can be grouped into five general categories as indicated below. Specific examples of each class are provided in **Table 1**.

1. **Cationic (basic) peptides.** These include polyarginines and peptoids ***(14)***, arginine-rich sequences ***(15)***, Tat ***(6)***, and penetratin ***(5)***.
2. **Hydrophobic sequences.** Including membrane translocating sequences (MTS) ***(16–18)***.
3. **Amphipathic sequences.** These present periodically hydrophobic and polar amino acids and include influenza virus hemagglutinin ***(19,20)*** and a range of synthetic peptides ***(13,21,22)***.

4. **Proline-rich *(23)* and antimicrobial sequences *(24,25)*.**
5. **Bipartite or chimeric peptides.** These constructs may contain two or more of the motifs listed. Examples include hydrophobic sequences from signal peptides with nuclear localization signal (NLS) ***(11,26)***, peptides derived from vascular endothelium cadherin (pVEC) ***(27)*** or the N-terminus of prion protein (PrP) ***(28)***, containing short flanking strongly hydrophobic and cationic motifs. The transportans contain both amphipathic and hydrophobic sequences ***(12)***. With some reservations, multifunctional branched peptides or loligomers ***(29,30)*** could also be classified in this group, although these are considered to be taken into cells by endocytosis.

Selection of an optimal CPP for delivering a specific cargo into a particular cell type, especially in vivo, is not necessarily a trivial task. Until now there are relatively few studies devoted to the comparison of the cellular delivery efficiency of different CPPs and the correlation to their side effects such as cytotoxicity and membrane destabilizing properties. Comparative data of the delivery efficiency of CPPs in vivo are even more scarce.

Comparison of four CPPs—penetratin ***(5)***, Tat ***(6)***, transportan ***(12)***, and model amphipathic peptide (MAP) ***(13)***—revealed that a model peptide cargo was most efficiently delivered into Bowes melanoma cells by MAP and transportan peptides. As judged by energy transfer experiments ***(31)***, the intracellular concentration of a cargo peptide delivered into cells by penetratin or Tat remained three- to fourfold lower compared with transportan- and MAP-mediated delivery. On the other hand, transportan and MAP were more noxious to cells and increased the plasma membrane permeability at lower concentrations. Import of penetratin sequences by the melanoma-derived SKMel-37 cells was in turn three- to fourfold more efficient than uptake of MTS-sequences as measured by fluorescence correlation spectroscopy in living cells and by FACS analysis ***(32)***.

Most CPPs are considered to display low toxicity and to be devoid of side effects in cultured cells. According to the MTT assay, the viability of different cell types in the presence of CPPs at 10 to 50 μ*M* concentrations has been reported to be excellent ***(33)***. However, the amphipathic CPPs, especially MAP, induce leakage of the plasma membrane at concentrations below 5 μ*M* ***(31)***. The toxicity of polyarginine peptides increases with the peptide length and polymers with molecular mass of 12 kDa are cytotoxic at concentrations less than 1 μ*M* ***(34)***.

Cell-penetrating peptides have delivered a wide variety of cargoes to cultured cells that range from small molecules like dyes and peptides to whole proteins ***(35–37)***, plasmids ***(38,39)***, phages ***(40)***, and superparamagnetic particles ***(41)***. Attachment of penetratin and Tat peptides to the surface of liposomes enhanced their cellular translocation efficiency in proportion to the number of attached peptides ***(42)***. Liposomes of 200 nm are the largest reported cargoes for which cellular uptake is induced by the Tat peptide and perhaps CPPs in general ***(43)***.

The cargoes of interest to be translocated into the cells differ vastly both in size and chemical properties. Thus, a range of different coupling strategies is required and both covalent and noncovalent coupling strategies, and combinations of both, have been utilized. In principle, noncovalent systems that simply rely on the ratiometric mixing of the cargo and CPP molecules would be clearly advantageous. Indeed, several types of cargoes have been successfully translocated by CPPs into cells as noncovalent complexes (**Fig. 1A–D**), taking advantage of the following interactions:

1. **Electrostatic interaction.** Cationic CPPs avidly form strong complexes with polyanions and induce the cellular uptake of oligonucleotides ***(44,45)*** and plasmids ***(30,46,47)***.
2. **Hydrogen bonds between specific pairs of nucleobases.** A short PNA oligomer covalently coupled to a CPP has been used for the cellular delivery of complementary oligonucleotides ***(48)*** and plasmids ***(39)***.
3. **Biotin–avidin interaction.** The cellular delivery capacity of several CPPs has been proven using complexes of biotinylated peptides with labeled avidin or streptavidin ***(23,27,37)***. Avidin could also serve as a scaffold for the assembly of molecules to be delivered into cells using biotinylated CPPs to facilitate the uptake process.
4. **Mixed-type interactions.** Short amphipathic carrier peptides like Pep-1, which wrap around proteins to induce their cellular uptake, were recently introduced ***(26)***. Interactions between Pep-1 and a protein probably involve both electrostatic and hydrophobic interactions together with hydrogen bonding. After passage across the plasma membrane the noncovalent complexes can dissociate in the cell interior, where the liberated cargo can find its cellular target.

However, connecting the cargo to a CPP by a covalent bond has been a more common strategy (**Fig. 1E–G**). The peptides and CPP are often assembled for cellular delivery as a continuous polypeptide chain by chemical synthesis ***(49)*** or by using recombinant DNA technology ***(50)***. The deliverable proteins are analogously produced in bacteria as fusion proteins with cell-penetrating sequences ***(51)***. Several expression vectors containing sequences coding for CPPs are also commercially available ***(52)***.

In parallel with stable chemical bonds, labile bonds have been used for CPP-cargo attachment. Especially popular has been the disulfide bond that is reduced in the cellular milieu ***(53)***, releasing cargo from the vector. The disulfide bond is formed between the CPP bearing an activated additional cysteine ***(44,54)*** and a thiol group of a cargo ***(55)***. Disulfide bond-containing constructs have been used for the delivery of oligonucleotides, peptides and proteins ***(56)***. Cargoes not usually possessing thiol groups are supplemented with these either by recombinant techniques or, more commonly, by chemical synthesis or modification.

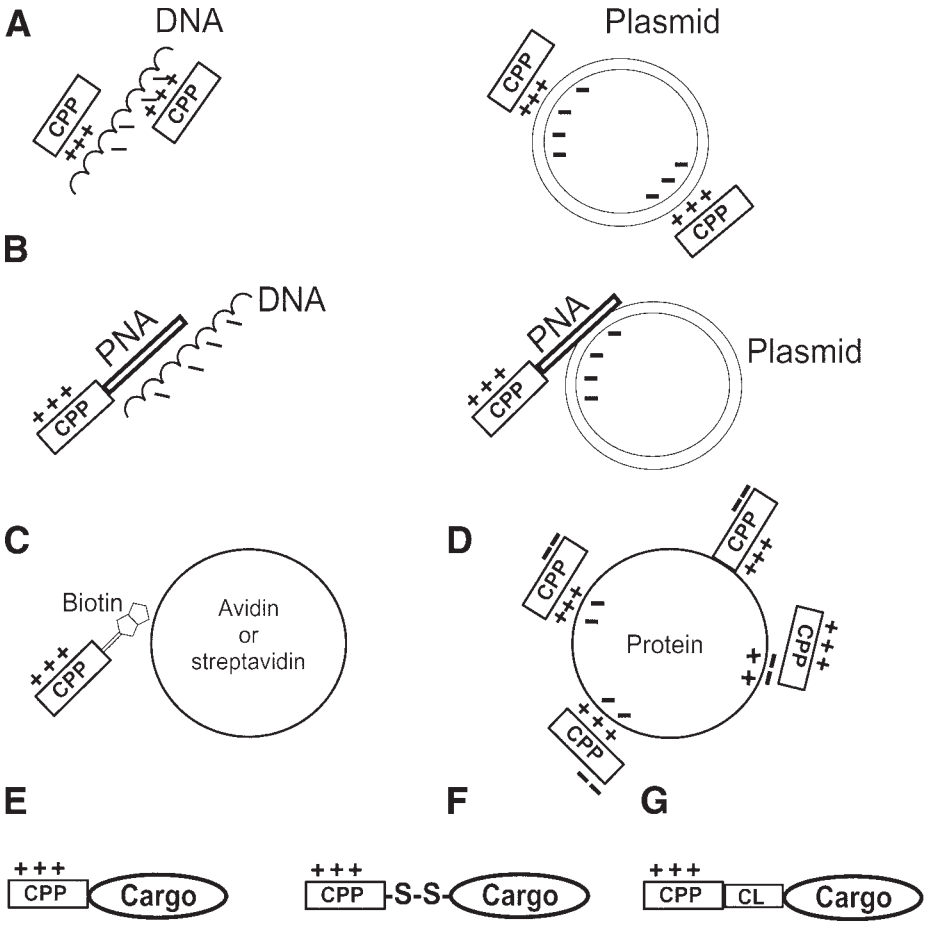


Fig. 1. Different strategies of coupling cargoes to cell-penetrating peptides. (I) Noncovalent coupling by electrostatic interactions **(A)** ***(44,46)***, specific pairing of nucleotides; **(B)** ***(39,48)***, biotin–avidin interaction; **(C)** ***(37,66)*** or mixed-type interactions **(D)** ***(26)***. (II) Covalent coupling by connecting CPP and cargo molecule into a continuous chain **(E)** ***(36)***, forming a disulfide bond **(F)** ***(49,56)*** or chemical crosslinking **(G)** ***(37,57)***.

A pioneering study of the cellular translocation of proteins demonstrated the uptake of enzymes that were connected to the Tat protein fragment by a chemical crosslink ***(57)***. Bifunctional crosslinkers can be used for more selective coupling of CPPs to proteins ***(37,58)***. A single-step ligation scheme by forming a thiazolidine ring ***(59)*** has been used for coupling the MTS sequence to a cargo peptide and several other orthogonal ligation strategies can be applied for the coupling of both protein and peptide cargoes to CPPs ***(60)***.

Interaction of CPPs with the cellular membrane, and their subsequent internalization, depends on the presence and location of specific motifs or amino acids in peptides, as shown most convincingly for W^6 of penetratin. Therefore, the position of cargo attachment to the CPP molecule needs careful consideration. Chemical synthesis enables the insertion of amino acids with activated or orthogonally protected side chain for cargo coupling in any desired position. Attachment of a fluorophore to the amino-terminus of penetratin or MTS yields efficiently internalizing peptides as measured by fluorescence correlation spectroscopy in living SKMel37 cells and by FACS analysis ***(32)***. Coupling of the same label to amino acids elsewhere in the sequence impaired the internalization efficiency to different degrees. Replacement of W^6 or F^7 by fluoresceinated lysine led to a dramatic five- to sixfold reduction of penetratin uptake but did not completely abolish cellular penetration. Substitutions in other positions of penetratin and MTS reduced their cellular uptake but not so profoundly ***(32)***. Thus, cargoes are preferentially coupled to an added extra cysteine at the N-terminal of penetratin ***(9)*** and the C-terminal of Tat ***(61)*** and MTS peptides ***(62)***. Transportan(s), in contrary, are preferably coupled to cargoes via the lysine side chain in the middle of the molecule ***(63)***, which is not a major determinant for membrane interaction ***(64)***. The properties of the deliverable peptide cargo, especially charge, greatly influence the cellular uptake efficiency ***(32)***. To conclude, the position of cargo attachment has to be considered for each particular CPP. Therefore, in order to prepare an efficiently internalizing construct, several features including the choice of CPP, a suitable coupling strategy, and the ratio between cargo and CPP have to be carefully considered.

2. Materials

2.1. Solid Phase Synthesis of Cell-Penetrating Peptides

2.1.1. Reagents

1. *t*-Butyloxycarbonyl (*t*-Boc) protected amino acids ***(65)***.
2. *p*-Methylbenzylhydrylamine (MBHA) resin, substitution about 1 mmol/g.
3. Hydroxybenzotriazole (HOBt) and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylammonium tetrafluoroborate (TBTU).
4. Kaiser test reagents:

Solution A: Mix solutions a and b.

a. 200 m*M* KCN in pyridine: dilute 2 mL of KCN solution (65 mg in 100 mL of water) in 100 mL of pyridine. KCN is highly toxic. KCN stock solution should be stored in the dark at 4°C.
b. Solution of 40 g phenol in 10 mL of dry ethanol. Phenol can cause severe burns. Phenol should be kept in the dark at 4°C.

Solution B: Ninhydrin solution.

Prepare 5% (w/v) ninhydrin solution in ethanol (500 mg of ninhydrin in 10 mL of ethanol).

5. Solvents and scavengers: methylene chloride (DCM), trifluoroacetic acid (TFA), *N*,*N*-diisopropylethylamine (DIEA), 100% ethanol, *N*,*N*-dimethylformamide, dimethylsulphoxide (DMSO), acetonitrile, trifluoromethanesulfonic acid (TFMSA), dimethylsulfide (DMS), *p*-cresol, dimercaptoethane.

2.1.2. Equipment

1. Applied Biosystems peptide synthesizer, model 431A.
2. Apparatus for HF cleavage.
3. Lyophilizer.
4. HPLC system with preparative and analytical C_{18} columns.
5. Mass spectrometer (MALDI TOF Voyager-DE STR, Applied Biosystems, USA).

3. Methods

3.1. Synthesis of Cell-Penetrating Peptides

Transport peptides can be synthesized using either *t*-Boc or Fmoc solid phase peptide synthesis strategies with a synthesizer or manually. We routinely synthesize CPPs in a stepwise manner on solid support using an Applied Biosystems Model 431A peptide synthesizer. *tert*-Butyloxycarbonyl amino acids are coupled as 1-hydroxybenzotriazole (HOBt) esters to a *p*-methylbenzylhydrylamine (MBHA) resin ***(65)***. C-terminally amidated CPPs are less prone to degradation and show higher internalization efficiency than carboxylic acid derivatives.

The N-terminal part of transportan inserts into SDS micelles and probably participates in the interaction with the plasma membrane of cells. Therefore, cargoes were coupled to ε-amino group of Lys^{14} residue in the middle of transportan. In this position Fmoc-protected Lys was introduced into the transportan sequence. The 3-nitro-2-pyridinesulphenyl (Npys) derivative of cysteine or biotin was coupled manually to the ε-amino group of this Lys residue either directly or via a glycine linker. Adding an extra Npys-Cys at N- or C-termini respectively modified penetratin and Tat. The cell-penetrating construct was obtained by forming an unsymmetrical disulfide between Cys-containing cargoes and (Npys)-Cys of CPP. Peptides containing S-Npys protected Cys have been shown to rapidly react with thiols to form disulphides at high yields ***(54)***. Here we present protocols for the coupling of (Npys)-Cys to CPPs and the cleavage and purification of respective peptides. As an example of producing a cell translocating construct a protocol for coupling a PNA oligomer to a CPP is also provided. Methods for the production of cell-penetrant peptides and protein–peptide conjugates can be found elsewhere ***(31,37)***.

3.1.1. Synthesis of Cys(Npys) Derivatives of Transport Peptides

1. Add DCM to 1 eq of protected peptide-resin, mix for 5 min, and drain.
2. Add 20% (v/v) piperidine/DMF to peptide-resin, mix 1 min, and drain.
3. Repeat step 2, mix for 10 min, and drain.
4. Wash five times with DMF and twice with DCM.
5. Perform Kaiser test to confirm the deprotection of Fmoc-protected Lys.
6. Dissolve 3 eq of Boc-Cys(Npys) in DCM, add 3 eq of HOBt and 3 eq of TBTU.
7. Add this solution to the peptide-resin and add 6 eq DIEA. The optimal volume of solution is approx 20 mL per gram of resin. Add DMF, if necessary.
8. Mix for 20 min.
9. Wash with DCM, ethanol, DMF, base (5% DIEA in DCM), DMF, three times with DCM. The optimal volume of each washing is about 20 mL per gram of resin, and all washes are carried out for 1 min.
10. Perform Kaiser test. If the resin stains blue, then the first coupling is not complete and recoupling is recommended (repeat **steps 6–10**).
11. After coupling the peptide-resin can be stored at 4°C (*see* **Note 1**).

3.1.2. Cleavage of Peptides From Resin and Purification

1. Final deprotection of the peptide-resin to remove *t*-Boc groups: Add approx 20 mL of 50% (v/v) TFA/DCM solution per gram resin, mix for 1 min, drain.
2. Repeat **step 1**, but mix for 20 min, drain.
3. Wash with DCM, ethanol, DMF, base, DMF, base, DMF, 3 times with DCM.
4. Perform Kaiser test.
5. Remove the formyl- and benzyl-type protective groups from peptide side chains by using the low TFMSA method. Treat the resin with a mixture of TFMSA/TFA/*p*-cresol/dimercaptoethane (10/50/30/8/2 v/v). Mix all the components on ice carefully adding TMFSA as the last step. Take 1 mL of mixture per 50 mg of peptide resin and mix on ice for 2 h.
6. Wash the deprotected resin with DCM, ethanol, DMF, base, DMF, base, three times with DCM.
7. Cleave the peptides from the MBHA resin using hydrogen fluoride according to the manufacturer's specifications (*see* **Note 2**).
8. Dissolve the crude peptide in 10% (v/v) acetic acid, remove the resin, and freeze-dry the peptide.
9. Purify the crude peptide by reverse phase HPLC using a preparative C_{18} column and freeze dry the fractions.
10. Confirm peptide mass using mass spectrometry.
11. Store freeze-dried peptides in the dark at 4°C (*see* **Note 1**).

3.1.3. Synthesis and Purification of Cell Penetrating Peptide-PNA Constructs

1. Dissolve for conjugation 1 eq of peptide and 1.5 eq of Cys-PNA in a deoxygenised mixture of DMSO/DMF/0.1 m*M* KH_2PO_4, pH 4.45 (v/v; 3:1:1) (*see* **Note 3**).

2. Continue the reaction overnight by stirring the solution in the dark under a nitrogen atmosphere (*see* **Note 1**).
3. Purify the resulting crude CPP-PNA construct by reversed phase HPLC on a semi-preparative C_{18} column. Elute constructs using a 50 min gradient (solvent A: 0.1% TFA in water; solvent B: 0.1% TFA in acetonitrile) from 20% to 100% solvent B. Record the absorbance at both 260 and 220 nm (*see* **Note 4**).
4. Freeze dry the construct.
5. Verify the product by MALDI-TOF mass spectrometry.
6. Store the construct at –20°C.

4. Notes

1. The Npys-group is light-sensitive and therefore the respective peptides are stored and reacted in the dark.
2. As a scavenger use only *p*-cresol because the Cys(Npys) residue is sensitive to *p*-thiocresol.
3. Organic solvents such as DMSO are used to increase the solubility of PNA.
4. The final yield of the construct was 35–65%. Yield depends on the length and sequence of PNA.

Acknowledgments

This work was supported by grants from the Swedish Institute Visby Program, Swedish Royal Academy of Sciences, Swedish Research Councils TFR and NFR, and CePeP Ltd. Sweden and Estonian Science Foundation (ESF 4007 and 5588).

References

1. Shen, W. C. and Ryser, H. J. (1978) Conjugation of poly-L-lysine to albumin and horseradish peroxidase: a novel method of enhancing the cellular uptake of proteins. *Proc. Natl. Acad. Sci. USA* **75,** 1872–1876.
2. Leonetti, J. P., Rayner, B., Lemaitre, M., et al. (1988) Antiviral activity of conjugates between poly(L-lysine) and synthetic oligodeoxyribonucleotides. *Gene* **72,** 323–332.
3. Letsinger, R. L., Zhang, G. R., Sun, D. K., Ikeuchi, T., and Sarin, P. S. (1989) Cholesteryl-conjugated oligonucleotides: synthesis, properties, and activity as inhibitors of replication of human immunodeficiency virus in cell culture. *Proc. Natl. Acad. Sci. USA* **86,** 6553–6556.
4. Langel, Ü., ed. (2002) *Cell-Penetrating Peptides, Processes and Applications*. CRC Press, Boca Raton, London, New York, Washington.
5. Derossi, D., Joliot, A. H., Chassaing, G., and Prochiantz, A. (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* **269,** 10,444–10,450.

6. Vivés, E., Brodin, P., and Lebleu, B. (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* **272,** 16,010–16,017.
7. Green, M. and Loewenstein, P. M. (1988) Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **55,** 1179–1188.
8. Frankel, A. D., Bredt, D. S., and Pabo, C. O. (1988) Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science* **240,** 70–73.
9. Derossi, D., Chassaing, G., and Prochiantz, A. (1998) Trojan peptides: the penetratin system for intracellular delivery. *Trends Cell Biol.* **8,** 84–87.
10. Lin, Y. Z., Yao, S. Y., Veach, R. A., Torgerson, T. R., and Hawiger, J. (1995) Inhibition of nuclear translocation of transcription factor NF-kappa B by a synthetic peptide containing a cell membrane-permeable motif and nuclear localization sequence. *J. Biol. Chem.* **270,** 14,255–14,258.
11. Chaloin, L., Vidal, P., Heitz, A., et al. (1997) Conformations of primary amphipathic carrier peptides in membrane mimicking environments. *Biochemistry* **36,** 11,179–11,187.
12. Pooga, M., Hällbrink, M., Zorko, M., and Langel, Ü. (1998) Cell penetration by transportan. *FASEB J.* **12,** 67–77.
13. Oehlke, J., Beyermann, M., Wiesner, B., et al. (1997) Evidence for extensive and non-specific translocation of oligopeptides across plasma membranes of mammalian cells. *Biochim. Biophys. Acta* **1330,** 50–60.
14. Wender, P. A., Mitchell, D. J., Pattabiraman, K., Pelkey, E. T., Steinman, L., and Rothbard, J. B. (2000) The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. *Proc. Natl. Acad. Sci. USA* **97,** 13,003–13,008.
15. Futaki, S., Suzuki, T., Ohashi, W., et al. (2001) Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. *J. Biol. Chem.* **276,** 5836–5840.
16. Lin, M., Nairn, A. C., and Guggino, S. E. (1992) cGMP-dependent protein kinase regulation of a chloride channel in T84 cells. *Am. J. Physiol.* **262,** C1304–1312.
17. Lin, Y. Z., Yao, S. Y., and Hawiger, J. (1996) Role of the nuclear localization sequence in fibroblast growth factor-1-stimulated mitogenic pathways. *J. Biol. Chem.* **271,** 5305–5308.
18. Liu, K. Y., Timmons, S., Lin, Y. Z., and Hawiger, J. (1996) Identification of a functionally important sequence in the cytoplasmic tail of integrin beta 3 by using cell-permeable peptide analogs. *Proc. Natl. Acad. Sci. USA* **93,** 11,819–11,824.
19. Lear, J. D. and DeGrado, W. F. (1987) Membrane binding and conformational properties of peptides representing the NH2 terminus of influenza HA-2. *J. Biol. Chem.* **262,** 6500–6505.
20. Plank, C., Zauner, W., and Wagner, E. (1998) Application of membrane-active peptides for drug and gene delivery across cellular membranes. *Adv. Drug Deliv. Rev.* **34,** 21–35.

21. Parente, R. A., Nadasdi, L., Subbarao, N. K., and Szoka, F. C. Jr. (1990) Association of a pH-sensitive peptide with membrane vesicles: role of amino acid sequence. *Biochemistry* **29,** 8713–8719.
22. Wyman, T. B., Nicol, F., Zelphati, O., Scaria, P. V., Plank, C., and Szoka, F. C. Jr. (1997) Design, synthesis, and characterization of a cationic peptide that binds to nucleic acids and permeabilizes bilayers. *Biochemistry* **36,** 3008–3017.
23. Sadler, K., Eom, K. D., Yang, J. L., Dimitrova, Y., and Tam, J. P. (2002) Translocating proline-rich peptides from the antimicrobial peptide bactenecin 7. *Biochemistry* **41,** 14,150–14,157.
24. Takeshima, K., Chikushi, A., Lee, K. K., Yonehara, S., and Matsuzaki, K. (2003) Translocation of analogues of the antimicrobial peptides magainin and buforin across human cell membranes. *J. Biol. Chem.* **278,** 1310–1315.
25. Kobayashi, S., Takeshima, K., Park, C. B., Kim, S. C., and Matsuzaki, K. (2000) Interactions of the novel antimicrobial peptide buforin 2 with lipid bilayers: proline as a translocation promoting factor. *Biochemistry* **39,** 8648–8654.
26. Morris, M. C., Depollier, J., Mery, J., Heitz, F., and Divita, G. (2001) A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat. Biotechnol*. **19,** 1173–1176.
27. Elmquist, A., Lindgren, M., Bartfai, T., and Langel, Ü. (2001) VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions. *Exp. Cell. Res.* **269,** 237–244.
28. Lundberg, P., Magzoub, M., Lindberg, M., et al. (2002) Cell membrane translocation of the N-terminal (1-28) part of the prion protein. *Biochem. Biophys. Res. Commun.* **299,** 85–90.
29. Sheldon, K., Liu, D., Ferguson, J., and Gariepy, J. (1995) Loligomers: design of de novo peptide-based intracellular vehicles. *Proc. Natl. Acad. Sci. USA* **92,** 2056–2060.
30. Singh, D., Bisland, S. K., Kawamura, K., and Gariepy, J. (1999) Peptide-based intracellular shuttle able to facilitate gene transfer in mammalian cells. *Bioconjug. Chem.* **10,** 745–754.
31. Hällbrink, M., Floren, A., Elmquist, A., Pooga, M., Bartfai, T., and Langel, Ü. (2001) Cargo delivery kinetics of cell-penetrating peptides. *Biochim. Biophys. Acta* **1515,** 101–109.
32. Fischer, R., Waizenegger, T., Kohler, K., and Brock, R. (2002) A quantitative validation of fluorophore-labelled cell-permeable peptide conjugates: fluorophore and cargo dependence of import. *Biochim. Biophys. Acta* **1564,** 365–374.
33. Pooga, M., Elmquist, A., and Langel, Ü. (2002) Toxicity and side effects of cell-penetrating peptides, in *Cell Penetrating Peptides, Processes and Applications* (Langel, Ü., ed.), CRC Press, Boca Raton, London, New York, Washington, pp. 245–261.
34. Mitchell, D. J., Kim, D. T., Steinman, L., Fathman, C. G., and Rothbard, J. B. (2000) Polyarginine enters cells more efficiently than other polycationic homopolymers. *J. Pept. Res.* **56,** 318–325.

35. Rojas, M., Donahue, J. P., Tan, Z., and Lin, Y. Z. (1998) Genetic engineering of proteins with cell membrane permeability. *Nat. Biotechnol.* **16,** 370–375.
36. Schwarze, S. R., Ho, A., Vocero-Akbani, A., and Dowdy, S. F. (1999) In vivo protein transduction: delivery of a biologically active protein into the mouse. *Science* **285,** 1569–1572.
37. Pooga, M., Kut, C., Kihlmark, M., et al. (2001) Cellular translocation of proteins by transportan. *FASEB J.* **15,** 1451–1453.
38. Morris, M. C., Chaloin, L., Heitz, F., and Divita, G. (2000) Translocating peptides and proteins and their use for gene delivery. *Curr. Opin. Biotechnol*. **1,** 461–466.
39. Branden, L. J., Mohamed, A. J., and Smith, C. I. (1999) A peptide nucleic acid-nuclear localization signal fusion that mediates nuclear transport of DNA. *Nat. Biotechnol.* **17,** 784–787.
40. Eguchi, A., Akuta, T., Okuyama, H., et al. (2001) Protein transduction domain of HIV-1 Tat protein promotes efficient delivery of DNA into mammalian cells. *J. Biol. Chem*. **276,** 26,204–26,210.
41. Lewin, M., Carlesso, N., Tung, C. H., et al. (2000) Tat peptide-derivatized magnetic nanoparticles allow in vivo tracking and recovery of progenitor cells. *Nat. Biotechnol*. **18,** 410–414.
42. Tseng, Y. L., Liu, J. J., and Hong, R. L. (2002) Translocation of liposomes into cancer cells by cell-penetrating peptides penetratin and tat: a kinetic and efficacy study. *Mol. Pharmacol.* **62,** 864–872.
43. Torchilin, V. P., Rammohan, R., Weissig, V., and Levchenko, T. S. (2001) Tat peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors. *Proc. Natl. Acad. Sci. USA* **98,** 8786–8791.
44. Oehlke, J., Birth, P., Klauschenz, E., et al. (2002) Cellular uptake of antisense oligonucleotides after complexing or conjugation with cell-penetrating model peptides. *Eur. J. Biochem.* **269,** 4025–4032.
45. Dom, J., Shaw-Jackson, C., Matis, C., et al. (2003) Cellular uptake of Antennapedia penetratin peptides is a two-step process in which phase transfer precedes a tryptophan-dependent translocation. *Nucl. Acids Res.* **31,** 556–561.
46. Morris, M. C., Chaloin, L., Mery, J., Heitz, F., and Divita, G. (1999) A novel potent strategy for gene delivery using a single peptide vector as a carrier. *Nucl. Acids Res.* **27,** 3510–3517.
47. Futaki, S., Ohashi, W., Suzuki, T., et al. (2001) Stearylated arginine-rich peptides: a new class of transfection systems. *Bioconjug. Chem.* **12,** 1005–1011.
48. Branden, L. J., Christensson, B., and Smith, C. I. (2001) In vivo nuclear delivery of oligonucleotides via hybridizing bifunctional peptides. *Gene Ther.* **8,** 84–87.
49. Lindgren, M., Hällbrink, M., Prochiantz, A., and Langel, Ü. (2000) Cell-penetrating peptides. *Trends Pharmacol. Sci.* **21,** 99–103.
50. Prochiantz, A. (1999) Homeodomain-derived peptides. In and out of the cells. *Ann. NY Acad. Sci.* **886,** 172–179.
51. Schwarze, S. R., Hruska, K. A., and Dowdy, S. F. (2000) Protein transduction: unrestricted delivery into all cells? *Trends Cell Biol.* **10,** 290–295.

52. www.qbiogene.com.
53. Feener, E. P., Shen, W. C., and Ryser, H. J. (1990) Cleavage of disulfide bonds in endocytosed macromolecules. A processing not associated with lysosomes or endosomes. *J. Biol. Chem.* **265,** 18,780–18,785.
54. Bernatowicz, M. S., Matsueda, R., and Matsueda, G. R. (1986) Preparation of Boc-[S-(3-nitro-2-pyridinesulfenyl)]-cysteine and its use for unsymmetrical disulfide bond formation. *Int. J. Pept. Protein Res.* **28,** 107–112.
55. Bernatowicz, M. S. and Matsueda, G. R. (1986) Preparation of peptide-protein immunogens using N-succinimidyl bromoacetate as a heterobifunctional crosslinking reagent. *Anal. Biochem.* **155,** 95–102.
56. Prochiantz, A. (2000) Messenger proteins: homeoproteins, TAT and others. *Curr. Opin. Cell Biol.* **12,** 400–406.
57. Fawell, S., Seery, J., Daikh, Y., et al. (1994) Tat-mediated delivery of heterologous proteins into cells. *Proc. Natl. Acad. Sci. USA* **91,** 664–668.
58. Pooga, M. and Langel, Ü. (2001) Targeting of cancer-related proteins with PNA oligomers. *Curr. Cancer Drug Targets* **1,** 231–239.
59. Zhang, L., Torgerson, T. R., Liu, X. Y., et al. (1998) Preparation of functionally active cell-permeable peptides by single-step ligation of two peptide modules. *Proc. Natl. Acad. Sci. USA* **95,** 9184–9189.
60. Tam, J. P., Yu, Q., and Miao, Z. (1999) Orthogonal ligation strategies for peptide and protein. *Biopolymers* **51,** 311–332.
61. Vivès, E., Dell'Aquila, C., Bologna, J. C., Morvan, F., Rayner, B., and Imbach, J. L. (1999) Lipophilic pro-oligonucleotides are rapidly and efficiently internalized in HeLa cells. *Nucl. Acids Res.* **27,** 4071–4076.
62. Chang, M., Zhang, L., Tam, J. P., and Sanders-Bush, E. (2000) Dissecting G protein-coupled receptor signaling pathways with membrane-permeable blocking peptides. Endogenous 5-HT(2C) receptors in choroid plexus epithelial cells. *J. Biol. Chem.* **275,** 7021–7029.
63. Pooga, M., Hällbrink, M., and Langel, Ü. (2002) Transportans, in *Cell Penetrating Peptides, Processes and Applications* (Langel, Ü., ed.), CRC Press, Boca Raton, London, New York, Washington, pp. 53–70.
64. Lindberg, M., Jarvet, J., Langel, Ü., and Gräslund, A. (2001) Secondary structure and position of the cell-penetrating peptide transportan in SDS micelles as determined by NMR. *Biochemistry* **40,** 3141–3149.
65. Soomets, U., Lindgren, M., Gallet, X., et al. (2000) Deletion analogues of transportan. *Biochim. Biophys. Acta* **1467,** 165–176.
66. Kilk, K., Magzoub, M., Pooga, M., Eriksson, L. E., Langel, Ü., and Gräslund, A. (2001) Cellular internalization of a cargo complex with a novel peptide derived from the third helix of the islet-1 homeodomain. Comparison with the penetratin peptide. *Bioconjug. Chem.* **12,** 911–916.

6

Incorporation of Phosphotyrosyl Mimetic 4-(Phosphonodifluoromethyl)phenylalanine (F_2Pmp) Into Signal Transduction-Directed Peptides

Zhu-Jun Yao, Kyeong Lee, and Terrence R. Burke, Jr.

Summary

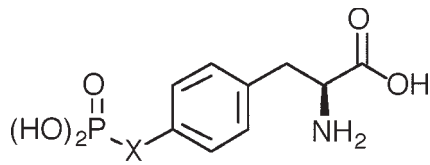
Phosphotyrosyl (pTyr)-containing sequences in proteins serve important roles in cellular signal transduction. Often, synthetic pTyr-containing peptides based on cognate sequences surrounding these pTyr residues can exhibit pharmacologically useful properties of full phosphoproteins. However, such pTyr-containing peptides have limited use in whole-cell systems resulting from lability of the phosphate ester bond to protein-tyrosine phosphatases (PTPs). For this reason, a number of phosphatase-stable pTyr mimetics have been developed that retain certain of the chemical and pharmacological properties of pTyr itself. Among these, difluoro-phosphonomethyl phenylalanine (F_2Pmp) has shown widespread utility in a variety of signal transduction settings, particularly those involving PTPs. This chapter provides practical techniques for the synthesis of a range of F_2Pmp-containing peptides.

Key Words: Signal transduction; pTyr mimetic; Pmp; F_2Pmp; protein-tyrosine kinase; protein-tyrosine phosphatase; SH2 domain; phosphopeptides.

1. Introduction

Signal transduction involves a transfer of information from cell surface receptors to the cytoplasm and ultimately to the nucleus where cellular activation and response are initated ***(1)***. Protein-tyrosine kinases (PTKs) serve central roles in many of these pathways by phosphorylating tyrosyl residues, which result in the introduction of phosphotyrosyl (pTyr, **1**, **Fig. 1**) pharmacophores

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

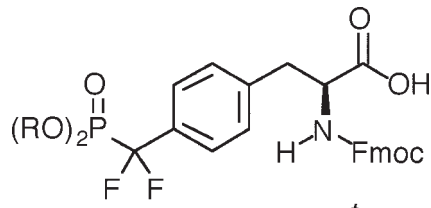


1 pTyr X = O
2 Pmp X = CH_2
3 F_2Pmp X = CF_2

Fig 1. Structures of pTyr and pTyr mimetics.

into cytoplasmic signaling proteins ***(2)***. Once formed, pTyr residues can alter enzyme activity in their own right or serve as recognition sites for binding by secondary signaling proteins by means of src homology 2 (SH2) or phosphotyrosyl binding (PTB) domains ***(3,4)***. The actions of PTKs can be modulated by protein-tyrosine phosphatases (PTPs) that dephosphorylate pTyr residues and thereby remove the pTyr motif ***(5)***. Because of the critical roles served by PTK-dependent signaling in normal cellular activation and homeostasis, disregulation of these pathways can contribute to a variety of pathogenic conditions, including several cancers ***(6)***. Modulation of PTK-dependent signaling by synthetic agents may potentially afford novel therapeutic approaches toward these diseases ***(7)***. The fact that pTyr residues provide critical features needed for substrate recognition by PTKs and PTPs as well as high affinity binding to SH2 domains, has made the pTyr pharmacophore an important component of inhibitor design ***(8–10)***.

The lability of the pTyr phosphate ester toward PTPs has necessitated development of pTyr mimetics that are stable to cellular phosphatases yet retain biophysical properties of parent pTyr residues. The compound phosphonomethyl phenylalanine (Pmp, 2), in which the phosphate ester linkage is replaced by a stable phosphonate methylene ***(11)***, is one analog that has shown wide utility as a phosphatase-stable pTyr mimetic in SH2 domain-directed inhibitors. Pmp-containing peptides can exhibit reduced SH2 domain-binding affinity relative to their pTyr counterparts. This has been attributed to higher pKa_2 values for phosphonate as compared with phosphate as well as to a loss of SH2 domain hydrogen bonding interactions at the phosphonate methylene that would normally be possible with the pTyr ester oxygen ***(12)***. Introduction of fluorines at this methylene to yield difluoro-phosphonomethyl phenylalanine (F_2Pmp, **3**) overcomes these limitations by lowering the phosphonate pKa_2 value to below that of parent pTyr and introducing potential hydrogen bonding between the CF_2 moiety and the SH2 domain ***(13)***.



4 R = Bu^t
5 R = Et
6 R = H

Fig. 2. Structures of *N*-Fmoc F_2Pmp derivatives.

The utility of F_2Pmp in SH2 domain-directed peptides varies among families of SH2 domains, with certain domains such as Src exhibiting enhanced affinity for F_2Pmp relative to Pmp, while domains such as Grb2 exhibit reduced affinity. In either case affinity differences are on the order of one- to fivefold ***(14)***. Alternatively, a much more dramatic effect is observed in PTP-directed peptides where F_2Pmp residues can exhibit 1000-fold enhanced affinity relative to Pmp ***(15)***. Therefore, F_2Pmp is an important amino acid derivative with wide utility in the synthesis of a variety of signal transduction-directed peptides, particularly for PTP-binding peptides ***(16)***.

The initial report of F_2Pmp was as its N-Fmoc-protected racemate bearing difluorophosphonate *tert*-butyl protection (**4**, **Fig. 2**) ***(17)***. Although removal of phosphonate protection can be achieved readily using TFA, tedious HPLC separation of diatereomeric D- and L-F_2Pmp-containing peptides is required. Alternatively, enantioselective syntheses of N-Fmoc (diethyl phosphono(difluoromethyl))-L-phenylalanine (*N*-Fmoc L-F_2Pmp$(OEt)_2$-OH, **5**) have been reported with the difluorophosphonate protected as its diethyl ester ***(18–20)***, which allows the preparation of diastereomerically pure peptides. However, harsh acidic conditions are required for subsequent removal of difluorophosphonate ethyl protection, which can result in unwanted side reactions and generation of heterogeneous products ***(21)***. To circumvent these difficulties, Gordeev et al. reported deprotection of *N*-Fmoc L-F_2Pmp$(OEt)_2$-OH (**5**) prior to peptide incorporation ***(22)***. Similar to *N*-Fmoc $((HO)_2PO)$-Tyr, which can be incorporated directly into peptides without phosphate protection ***(23)***, the resulting *N*-Fmoc (phosphono (difluoromethyl))-L-phenylalanine (*N*-Fmoc L-F_2Pmp-OH, **6**) can be utilized directly for peptide synthesis, obviating the need to subject peptides to harsh deprotection conditions. The importance of F_2Pmp in signal transduction studies renders the preparation of F_2Pmp-containing peptides of widespread interest. For this reason, this chapter is intended to provide the peptide chemist with a practical guide for synthesizing F_2Pmp-containing peptides.

2. Materials

1. *N*-Fmoc (diethyl phosphono(diflouromethyl))-L-phenylalanine (*N*-Fmoc L-F_2Pmp $(OEt)_2$-OH, 5) (CAS registry no. 160751-44-0) can be purchased from ChemPacific Inc. (Baltimore, MD) (*see* **Notes 1** and **2**).
2. 4-(2',4'-Dimethoxyphenyl-Fmoc-aminomethyl)-phenoxymethyl-polystyrene resin (*N*-Fmoc Rink amide resin ***[24]***) can be purchased from a variety of sources. Work reported herein used resin purchased from Bachem Bioscience Inc. (King of Prussia, PA) (*see* **Note 3**).
3. Vydac preparative reverse-phase Protein & Peptide C_{18} HPLC column (*see* **Note 4**).
4. Other *N*-Fmoc-amino acids: Fmoc-L-Asp(^{t}Bu)-OH; *N*-Fmoc-L-Asn-OH; *N*-Fmoc-L-Glu(^{t}Bu)-OH; *N*-Fmoc-Leu-OH; *N*-Fmoc-L-Ser(^{t}Bu)-OH; *N*-Fmoc-L-Thr(^{t}Bu)-OH; *N*-Fmoc L-Val-OH; *N*-*N*-Fmoc L-Ile-OH.
5. N-Methyl-2-pyrrolidinone (NMP).
6. *Bis*(trimethylsilyl)trifluoroacetamide (BSTFA).
7. Anhydrous CH_2Cl_2.
8. Me_3SiI.
9. Trifluoroacetic acid (TFA).
10. Acetonitrile (MeCN).
11. 1-Hydroxybenzotriazole hydrate (HOBt).
12. *N,N′*-diisopropylcarbodiimide (DIPCDI).
13. Piperidine.
14. Benzotriazole-1-yloxy-tris(dimethylamino)-phosphonium hexafluorophosphate (BOP).
15. Diisopropylethylamine (DIPEA).
16. Anisole.
17. Triethylsilane (TES).
18. 1-Acetylimidazole.
19. 4-Dimethylaminopyridine (DMAP).

3. Methods

3.1. Conversion of Commercially Available N*-Fmoc (Diethyl Phosphono(diflouromethyl))-L-Phenylalanine (5) to* N*-Fmoc (Phosphono(diflouromethyl))-L-Phenylalanine (*N*-Fmoc-L-F_2Pmp-OH* **[6]***)*

At the time of this writing, *N*-Fmoc (phosphono(diflouromethyl))-L-phenylalanine (**6**) is not commercially available. However, it can be prepared from commercially available *N*-Fmoc (diethyl phosphono(diflouromethyl))-L-phenylalanine (**5**) (*see* **Notes 1** and **2**) as outlined in the following steps.

1. According to the method of Gordeev et al. ***(22)***, a mixture of 1.76 g (3.07 mmol) *N*-Fmoc (diethyl phosphono(diflouromethyl))-L-phenylalanine **(5)** and 8.98 mL (8.70 g, 33.8 mmol) bis(trimethylsilyl)trifluoroacetamide (BSTFA) in 35 mL anhydrous CH_2Cl_2 was stirred under argon for 1 h at room temperature.

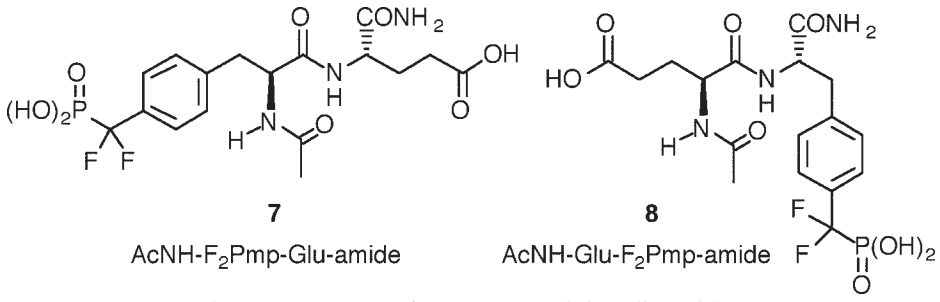


Fig. 3. Structures of F_2Pmp-containing dipeptides.

2. The mixture was cooled to –20°C, 3.50 mL (4.91 g, 24.5 mmol) Me_3SiI was added, and the mixture was allowed to warm to room temperature with continued stirring for 3 h.
3. Solvent was removed by rotary evaporation under reduced pressure and excess volatiles were further removed under high vacuum overnight to provide 2.9 of crude product as a thick brown oil.
4. The oil was treated at room temperature with a mixture of 5 mL TFA: 5 mL H_2O: 10 mL MeCN for 1 h. Solvent was removed by rotary evaporation and the residue was placed under high vacuum to yield 1.56 g of product as a foam that was pulverized to a pale yellow solid (98% yield). (Purity was >90% as indicated by H NMR; material was stored at –25°C.) H NMR (DMSO-d_6) δ: 8.77 (1H, brs); 8.49 (1H, brs); 7.88 (2H, δ, J = 7.4 Hz); 7.78 (1H, d, J = 8.5 Hz); 7.68 (2H, t, J = 6.8 Hz); 7.39 (8H, m); 4.89 (brs); 4.22 (7H, m); 3.45 (1H, ddd, J = 13.8 Hz & 4.5 Hz & 1.8 Hz); 3.13 (1H, dd, J = 13.6 Hz & 3.7 Hz); 2.93 (1H, dd, J = 10.7 & 13.6 Hz).

3.2. Incorporation of N*-Fmoc (Phosphono(diflouromethyl))-*L*-Phenylalanine Into Peptides*

*3.2.1. F_2Pmp-Containing Dipeptides (**Fig. 3**)*

3.2.1.1. AcNH-F_2Pmp-Glu-Amide *(7)*

1. *N*-Fmoc Rink amide resin ***(24)*** (110 mg, 0.05 mmol) (*see* **Note 3**) was pre-swollen with *N*-methyl-2-pyrrolidinone (NMP) (5 × 2 mL) and then deblocked by treatment with 20% (v/v) piperidine in 1 mL NMP at room temperature for 20 min.
2. The resin was washed with NMP (5 × 1 mL), and then coupled with the active ester of *N*-Fmoc-L-Glu(^{t}Bu)-OH, which was freshly prepared in a separate vessel by mixing 65 mg (0.15 mmol) *N*-Fmoc-L-Glu(^{t}Bu)-OH, 21 mg (0.15 mmol) 1-hydroxybenzotriazole hydrate (HOBt), 24 µL (0.15 mmol) *N*,*N*'-diisoproplcarbodiimide (DIPCDI) in 1 mL NMP at room temperature for 20 min, and then adding to the resin. Agitation (by gentle rocking motion) was continued for 2 h, then the reaction solution was removed by filtration and the resin was washed with NMP (5 × 1 mL).

3. Deprotection by treatment with 20% (v/v) piperidine in 1 mL NMP was conducted as described in **step 1** and the resin was washed with NMP (5 × 1 mL).
4. The resin was then treated with the active ester of *N*-Fmoc-L-F_2Pmp-OH (***6***), which was prepared in a separate vessel by stirring a mixture of 103 mg (0.20 mmol) *N*-Fmoc-L-F_2Pmp-OH, 28 mg (0.20 mmol) HOBt, 88 mg (0.20 mmol) benzotriazole-1-yloxy-tris(dimethylamino)-phosphonium hexafluorophosphate (BOP), 70 µL (0.40 mmol) diisopropylethylamine (DIPEA) in 1 mL NMP at room temperature for 10 min prior to transferring to the resin.
5. The resin was filtered and washed with NMP (5 × 1 mL) then treated again with 20% (v/v) piperidine in 1 mL NMP for 20 min and washed with NMP (5 × 1 mL).
6. The terminal free amino group was acetylated by treatment with 28 mg (0.25 mmol) 1-acetylimidazole in 1 mL NMP overnight.
7. After washing with NMP (5 × 1 mL) and CH_2Cl_2 (5 × 1 mL) the dipeptide was cleaved from the resin using a mixture of 2 mL TFA-anisole 95:5) for 30 min. The filtrate was collected and the resin was washed with CH_2Cl_2:TFA (1:1) (10 × 1 mL) and the combined washes were taken to dryness under vacuum.
8. The residue was dissolved in 4 mL of H_2O:MeCN (1:1) and filtered through a 0.45-µ filter. Preparative reverse-phase HPLC was conducted using a Vydac Protein & Peptide C_{18} column (*see* **Note 4**) with a flow rate of 20 mL/min; UV detection was at 254 nm. Solvent conditions were: B = 0% from 0 to 5 min with a linear gradient from 0% B to 50% B over 20 min. Solvents were: A = 0.05% TFA in H_2O and B = 0.05% TFA in MeCN. Product was collected at 11.3 min and lyophilized to provide **7** as a white solid (7.8 mg, 34% based on resin substitution). H NMR (D_2O, 250 MHz) δ: 7.57 (2H, d, *J* = 7.9 Hz), 7.37 (2H, d, *J* = 8.1 Hz), 4.58 (1H, t, *J* = 7.9 Hz), 4.27 (1H, dd, *J* = 5.1 Hz & 9.5 Hz), 3.12 (2H, d, *J* = 7.7 Hz), 2.36 (2H, t, *J* = 6.4 Hz), 2.05 (1H, m), 1.99 (3H, s), 1.88 (1H, m). FABMS (-Ve) *m/z*: 464 [(M-H)-, 100.0], 335 [(M-H-Glu)-, 15].

3.2.1.2. AcNH-Glu-F_2Pmp-Amide ***(8)***

1. *N*-Fmoc Rink amide resin (110 mg, 0.05 mmol) was pre-swollen with NMP and piperidine-deblocked as described in **Subheading 3.2.1.1., step 1**.
2. The resin was then double-coupled with the active ester of *N*-Fmoc-L-F_2Pmp-OH, which was prepared in a separate vessel as described in **Subheading 3.2.1.1., step 4** by stirring a mixture of 78 mg (0.15 mmol) *N*-Fmoc-L-F_2Pmp-OH (**6**), 21 mg (0.15 mmol) HOBt, 66 mg (0.15 mmol) BOP, and 52 µL (0.30 mmol) DIPEA in 1 mL NMP at room temperature for 10 min.
3. The resin was filtered and washed with NMP (5 × 1 mL), and then deblocked with 20% piperidine in 1 mL NMP.
4. After being washed with NMP (5 × 1 mL), the resin was double-coupled with the active ester of *N*-Fmoc-L-Glu(^{t}Bu)-OH, which was freshly prepared by mixing 108 mg (0.25 mmol) *N*-Fmoc-L-Glu(^{t}Bu)-OH, 35 mg (0.25 mmol) HOBt, 110 (0.25 mmol) BOP, and 87 µL (0.50 mmol) DIPEA in 1 mL NMP at room temperature for 10 min.

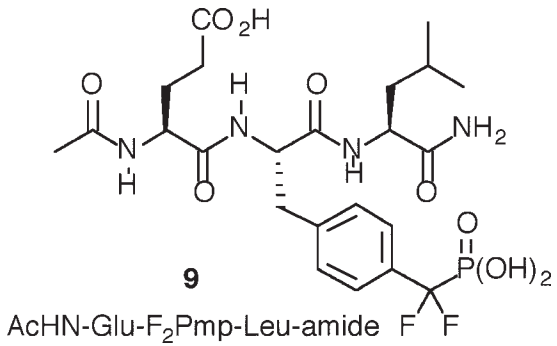


Fig. 4. F_2Pmp-containing tripeptide structure.

5. The resin was filtered and washed with NMP (5 × 1 mL) then treated for 20 min with 20% (v/v) piperidine in 1 mL NMP and washed with NMP (5 × 1 mL).
6. The terminal free amino group was acetylated using 28 mg (0.25 mmol) 1-acetylimidazole, 35 mg (0.25 mmol) HOBt, 110 (0.25 mmol) BOP, and 87 µL (0.50 mmol) DIPEA in 1 mL NMP for 2 h.
7. The resin was then washed with NMP (5 × 1 mL) and CH_2Cl_2 (5 × 1 mL), then the dipeptide was cleaved from the resin using a 2-mL mixture of TFA-H_2O-triethylsilane (TES) (92.5:5:2.5) for 1 h.
8. The filtrate was collected and combined with 10 × 1 mL washes (CH_2Cl_2:TFA, 1:1) and the total was concentrated to dryness. The residue was purified by HPLC as described in **Subheading 3.2.1.1., step 8** using a gradient of B = 0% from 0 to 5 min, then from 0% to 50% over 20 min. Product was collected at 11.1 min and lyophilized to provide 8 in its mono diisopropylethylamine salt form as a white solid (9.0 mg, 30% yield based on resin substitution). H NMR (D_2O, 250 MHz) δ: 7.55 (2H, d, *J* = 7.9 Hz), 7.38 (2H, d, *J* = 8.0 Hz), 4.70 (1H, dd, *J* = 5.6 Hz & 9.5 Hz), 4.21 (1H, t, *J* = 7.3 Hz), 3.73 (2H, m), 3.31 (1H, dd, *J* = 13.8 Hz & 5.7 Hz), 3.22 (2H, q, *J* = 7.5 Hz), 3.02 (1H, 1H, dd, *J* = 9.8 Hz & 13.9 Hz), 2.26 (2H, m), 2.00 (3H, s), 1.86 (2H, m), 1.35 (15H, m) ppm. FABMS (-Ve) *m/z*: 464 [(M-H)-, 20.5].

3.2.2. F_2Pmp-Containing Tripeptide

3.2.2.1. AcHN-Glu-F_2Pmp-Leu-Amide *(9)* (**Fig. 4**)

1. *N*-Fmoc Rink amide resin (220 mg, 0.10 mmol) was deblocked and washed as described earlier.
2. Coupling of amino acid residues were as their HOBt active esters, which were prepared as in **Subheading 3.2.1.2., step 4** with the exception except that 5 equivalents of *N*-Fmoc-L-Leu-OH, *N*-Fmoc-L-F_2Pmp-OH and *N*-Fmoc-L Glu(^{t}Bu)-OH were used in that order.

3. The terminal free amino group was acetylated using 110 mg (1.00 mmol) 1-acetylimidazole in 1 mL NMP and the peptide was cleaved from the resin using a 2-mL mixture of TFA:H_2O:triethylsilane (TES) (92.5:5:2.5) for 1 h.
4. The crude peptide was precipitated using cold ether and purified by HPLC as in **Subheading 3.2.1.1., step 8** using a gradient of B = 0% from 0 to 5 min, then from 0% to 50% over 20 min. Product was collected at 15.4 min and lyophilized to provide 9 as a white solid (25.4 mg, 30% yield based on resin substitution). NMR (D_2O, 250 MHz) δ: 7.58 (2H, d, *J* = 8.2 Hz), 7.49 (2H, d, *J* = 8.2 Hz), 4.71 (1H, t, *J* = 7.9 Hz), 4.21 (2H, m), 3.13 (2H, m), 2.34 (2H, m), 2.03 (3H, s), 1.96 (2H, m), 1.57 (3H, m), 0.92 (3H, d, *J* = 5.6 Hz), 0.86 (3H, d, *J* = 5.1 Hz). FABMS (-Ve) *m/z*: 577 [(M-H)-, 100.0].

3.2.3. F_2Pmp-Containing Hexapeptides

3.2.3.1. AcHN-Asp-Ala-Asp-Glu-F_2Pmp-Leu-Amide ***(10)*** (**Fig. 5**)

1. Rink amide resin (220 mg, 0.10 mmol) was pre-swollen with NMP (5 × 2 mL) and then deblocked using 20% (v/v) piperidine in 1 mL NMP at room temperature for 20 min.
2. The resin was then washed with NMP (5 × 1 mL) and coupled with the active ester of *N*-Fmoc-L-Leu-OH, which was freshly prepared by mixing 177 mg (0.50 mmol) *N*-Fmoc-L-Leu-OH, 68 mg (0.50 mmol) HOBt, and 79 mL (0.50 mmol) DIPCDI in 1 mL NMP for 20 min at room temperature.
3. After being rocked for 2 h, the resin was filtered and washed with NMP (5 × 2 mL) and deblocked with 20% (v/v) piperidine in 1 mL NMP for 20 min.
4. The resin was washed with NMP (5 × 2 mL) and then rocked for 2 h with the active ester of *N*-Fmoc-L-F_2Pmp-OH (5 eq), which was prepared as in **Subheading 3.2.1.1., step 4**. Coupling was done in the presence of 6 mg 4-dimethylaminopyridine (DMAP) (*see* **Notes 5** and **6**).
5. The resin was filtered and washed with NMP (5 × 2 mL), treated and deblocked with 20% (v/v) piperidine in 1 mL NMP for 20 min, then washed with NMP (5 × 2 mL).
6. Subsequent couplings of residues were accomplished using *N*-Fmoc L-Glu(^{t}Bu)-OH, *N*-Fmoc L-Asp(^{t}Bu)-OH, *N*-Fmoc L-Ala-OH and *N*-Fmoc-L-Asp(^{t}Bu)-OH in that order (*see* **Note 7**). Residues were converted to their HOBT active esters (5 eq) as in **Subheading 3.2.1.1., step 2** and coupled for 2 h each.
7. Following Fmoc-deprotection of the final residue, the terminal free amino group was acetylated using 110 mg (1.00 mmol) 1-acetylimidazole in 1 mL NMP overnight.
8. The resin was washed with NMP (5 × 2 mL) and CH_2Cl_2 (5 × 2 mL) and the peptide was cleaved from the resin using a 2-mL mixture of TFA:H_2O:TES (92.5:5:2.5) for 1 h.
9. The resin was filtered and washed with CH_2Cl_2:TFA (1:1) (10 × 1 mL) and the combined filtrates were concentrated under vacuum and crude peptide was precipitated using 20 mL cold ether. The resulting solid was washed with cold ether and dissolved in 8 mL of H_2O:MeCN (1:1).

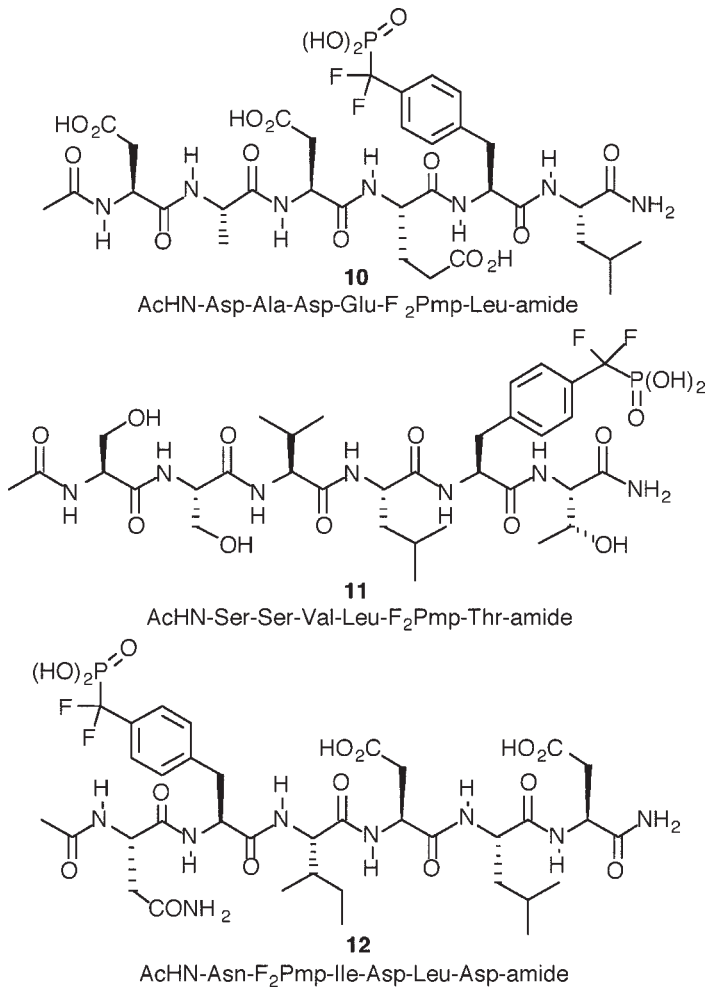


Fig. 5. Structures of F_2Pmp-containing hexapeptides.

10. The resulting solution was filtered through a 0.45-µm filter and acidified using TFA (50 µL). Purification by HPLC was performed as in **Subheading 3.2.1.1., step 8** using a gradient of B = 0% from 0 to 5 min then 0% to 50% over 25 min. Fractions were collected at 16.6 min and lyophilized to provide 10 as a white solid (45.7 mg, 52% yield based on resin substitution). H NMR (D_2O, 250 MHz): 7.56 (2H, d, J = 8.4 Hz), 7.38 (2H, d, J = 8.4 Hz), 4.66 (3H, m), 4.28 (3H, m), 3.17 (2H, m), 2.93 (4H, m), 2.49 (1H, m), 2.31 (1H, m), 2.04 (3H, s), 2.00 (2H, m), 1.64 (1H, m), 1.56 (2H, m), 1.41 (3H, d, J = 7.3 Hz), 0.91 (3H, d, J = 6.0 Hz), 0.84 (3H, d, J = 6.0 Hz). FABMS (-Ve) *m/z*: 878 [(M-H)-, 100.0]. Amino acid analysis: Asp 1.99 (2), Glu 1.00 (1), Ala 0.85 (1), Leu 1.05 (1).

3.2.3.2. AcHN-Ser-Ser-Val-Leu-F_2Pmp-Thr-Amide *(11)*

1. *N*-Fmoc Rink amide resin (220 mg, 0.10 mmol) was washed and deblocked as in **Subheading 3.2.1.1., step 1**.
2. The coupling of amino acid residues were accomplished as their HOBt active esters as in **Subheading 3.2.1.1., step 2**, using 5 eq relative to resin substitution. Amino acids were coupled in the order: *N*-Fmoc L-Thr(^{t}Bu)-OH, *N*-Fmoc L-F_2Pmp-OH, *N*-Fmoc L-Leu-OH, *N*-Fmoc L-Val-OH, *N*-Fmoc L-Ser(^{t}Bu)-OH and *N*-Fmoc L-Ser(^{t}Bu)-OH.
3. The terminal free amino group was acetylated with 1-acetylimidazole (110 mg, 1.00 mmol) in 1 mL NMP.
4. Peptide product was cleaved from the resin as in **Subheading 3.2.1.2., step 7** and precipitated using cold ether.
5. Purification by HPLC was performed as in **Subheading 3.2.1.1., step 8** using a gradient of B = 0% from 0 to 5 min then 0% to 50% over 25 min. Fractions were collected at 16.6 min and lyophilized to provide 11 as a white solid (33.6 mg, 41% yield based on resin substitution). H NMR (D_2O, 250 MHz): 7.56 (2H, d, J = 8.0 Hz), 7.37 (2H, d, J = 8.0 Hz), 4.73 (1H, m), 4.46 (2H, m), 4.30 (1H, m), 4.22 (1H, m), 4.12 (1H, m), 3.86 (5H, m), 3.16 (2H, d, J = 7.5 Hz), 2.06 (3H, s), 1.56 (4H, m), 1.17 (3H, dd, J = 6.0, 15.0 Hz), 0.92 (12H, m). FABMS *m/z*: 822.7 [(M-H)-, 100.0]. Amino acid analysis: Ser 1.91 (2), Thr 1.07 (1), Val 1.06 (1), Leu 1.00 (1).

3.2.3.3. AcHN-Asn-F_2Pmp-Ile-Asp-Leu-Asp-Amide *(12)*

1. *N*-Fmoc Rink amide resin (220 mg, 0.10 mmol) was washed and deblocked as in **Subheading 3.2.1.1., step 1**.
2. The coupling of amino acid residues was accomplished as their HOBt active esters as in **Subheading 3.2.1.1., step 2**, using 5 eq relative to resin substitution. Amino acids were coupled in the order: *N*-Fmoc-L-Asp(^{t}Bu)-OH, *N*-Fmoc L-Leu-OH, *N*-Fmoc L-Asp(^{t}Bu)-OH, *N*-Fmoc L-Ile-OH, *N*-Fmoc L-F_2Pmp-OH, and *N*-Fmoc-L-Asn-OH.
3. The terminal free amino group was acetylated with 1-acetylimidazole (110 mg, 1.00 mmol) in 1 mL NMP.
4. Peptide product was cleaved from the resin as in **Subheading 3.2.1.2., step 7** and precipitated using cold ether. Purification by HPLC was performed as in **Subheading 3.2.1.1., step 8** using a gradient of B = 0% from 0 to 5 min then 0% to 50% over 25 min. Fractions were collected at 18.9 min and lyophilized to provide 12 as a white solid (35.2 mg, 39% yield based on resin substitution). H NMR (D_2O, 250 MHz) δ: 7.57 (2H, d, J = 8.1 Hz), 7.35 (2H, d, J = 8.1 Hz), 4.72 (3H, m), 4.62 (1H, m), 4.33 (1H, m), 4.12 (1H, m), 3.26 (1H, dd), 2.75 approx 3.15 (5H, m), 2.67 (1H, dd, J = 5.9, 16.2 Hz), 2.53 (1H, dd, J = 8.2, 16.2 Hz), 1.97 (3H, s), 1.80 (1H, m), 1.62 (3H, m), 1.42 (1H, m), 1.14 (1H, m), 0.80 approx 0.97 (12H, m). FABMS *m/z*: 905 [(M-H)-, 100]. Amino acid analysis: Asn and Asp 3.00 (3), Ile 0.97 (1), Leu 1.02 (1).

4. Notes

1. ChemPacific USA Sales Marketing and Research Center, 6200 Freeport Centre, Baltimore, MD 21224, USA. Phone: 410-633-5771; Fax: 410-633-5808; Email: sales@chempacific.com; Website: http://www.chempacific.com.
2. *N*-Fmoc L-F_2Pmp$(OEt)_2$-OH **(5)** is indexed within the ACS system under registry number 160751-44-0 bearing the chemical name, 4-(difluoro-di-ethyl-phosphono-methyl)-*N*-[(9H-9-ylmethoxy)carbonyl]-L-phenylalanine. The associated chemical structure is currently incorrectly given as the free phosphonic acid, *N*-Fmoc L-F_2Pmp-OH (**6**).
3. Bachem Bioscience Inc., 3700 Horizon Drive, King of Prussia, PA 19406, USA. Phone: 800-634-3183; Fax: 610-239-0800; Website: www.bachem.com.
4. Catalogue no. 218TP1022 (22 mm diameter × 250 mm length; 10 μm pore size; 300 Å particle size); Website: http://www.vydac.com.
5. Coupling of *N*-Fmoc-F_2Pmp-OH can be difficult and may not go to completion. Following one standard coupling cycle, the resin should be checked using the Kaiser test, which provides a colorimetric indicator of unreacted terminal amines ***(25)***. A step-by-step protocol for performing the Kaiser test is available ***(26)*** and alternatives are reported in this book (*see* Chapters 1 and 5).
6. If the initial standard coupling of *N*-Fmoc-F_2Pmp-OH (5 eq for 2 h) fails to go to completion as indicated by the Kaiser test, the resin should be washed and a repeat coupling run overnight.
7. Coupling of the residue *N*-proximal to the F_2Pmp residue may occur sluggishly. If the initial standard HOBt active ester coupling of this residue using 5 eq each of *N*-Fmoc amino acid, HOBt, and DIPCDI in NMP for 20 min at room temperature prior to adding to the resin fails to provide a negative Kaiser test, then coupling should be repeated using an active ester solution formed by reacting 5 eq each of *N*-Fmoc amino acid, HOBT, and BOP with 10 eq DIPEA and 0.5 eq of DMAP 20 min at room temperature prior to adding to the resin.

Acknowledgments

Appreciation is expressed to Dr. James Kelley of the LMC for mass spectral analysis.

References

1. Edwards, D. R. (1994) Cell signalling and the control of gene transcription. *Trends Pharmacol. Sci.* **15,** 239–244.
2. Hunter, T. (2000) Signaling—2000 and beyond. *Cell* **100,** 113–127.
3. Bradshaw, J. M. and Waksman, G. (2003) Molecular recognition by SH2 domains. *Advances in Protein Chemistry* **61,** 161–210.
4. Borg, J. P. and Margolis, B. (1998) Function of PTB domains. *Protein Modules in Signal Transduction* **228,** 23–38.

5. Denu, J. M. and Dixon, J. E. (1998) Protein tyrosine phosphatases: mechanisms of catalysis and regulation. *Curr. Opin. Chem. Biol.* **2,** 633–641.
6. Bridges, A. (1995) The emerging role of protein phosphorylation and cell cycle control in tumor progression. *Chemtracts Organic Chemistry* **8,** 73–107.
7. Cohen, P. (1999) The development and therapeutic potential of protein kinase inhibitors. *Curr. Opin. Chem. Biol.* **3,** 459–465.
8. Burke, T. R. Jr., Yao, Z.-J., Smyth, M. S., and Ye, B. (1997) Phosphotyrosyl-based motifs in the structure-based design of protein-tyrosine kinase-dependent signal transduction inhibitors. *Curr. Pharm. Des.* **3,** 291–304.
9. Burke, T. R. Jr., Gao, Y., and Yao, Z.-J. (2000) Phosphoryltyrosyl mimetics as signaling modulators and potential antitumor agents. In *Biomedical Chemistry: Applying Chemical Principles to the Understanding and Treatment of Disease,* 1st ed. (Torrence, P. F., ed.), John Wiley & Sons, New York, pp. 189–210.
10. Burke, T. R. Jr., Yao, Z.-J., Liu, D.-G., Voigt, J., and Gao, Y. (2001) Phosphoryltyrosyl mimetics in the design of peptide-based signal transduction inhibitors. *Biopolymers* **60,** 32–44.
11. Marseigne, I. and Roques, B. P. (1988) Synthesis of new amino acids mimicking sulfated and phosphorylated tyrosine residues. *J. Org. Chem.* **53,** 3621–3624.
12. Domchek, S. M., Auger, K. R., Chatterjee, S., Burke, T. R., and Shoelson, S. E. (1992) Inhibition of SH2 domain/phosphoprotein association by a nonhydrolyzable phosphonopeptide. *Biochem.* **31,** 9865–9870.
13. Smyth, M. S., Ford, H. Jr., and Burke, T. R. Jr. (1992) A general method for the preparation of benzylic alpha, alpha-difluorophosphonic acids; non-hydrolyzable mimetics of phosphotyrosine. *Tetrahedron Lett.* **33,** 4137–4140.
14. Burke, T. R. Jr., Smyth, M. S., Otaka, A., et al. (1994) Nonhydrolyzable phosphotyrosyl mimetics for the preparation of phosphatase-resistant SH2 domain inhibitors. *Biochemistry* **33,** 6490–6494.
15. Burke, T. R., Kole, H. K., and Roller, P. P. (1994) Potent inhibition of insulin receptor dephosphorylation by a hexamer peptide containing the phosphotyrosyl mimetic F(2)Pmp. *Biochem. Biophys. Res. Commun.* **204,** 129–134.
16. Burke, T. R. and Zhang, Z. Y. (1998) Protein-tyrosine phosphatases: Structure, mechanism, and inhibitor discovery. *Biopolymers* **47,** 225–241.
17. Burke, T. R., Smyth, M. S., Nomizu, M., Otaka, A., and Roller, P. P. (1993) Preparation of fluoro-4-(phosphonomethyl)-D,L-phenylalanine and hydroxy-4-(phosphonomethyl)-D,L-phenylalanine suitably protected for solid-phase synthesis of peptides containing hydrolytically stable analogs of O-phosphotyrosine. *J. Org. Chem.* **58,** 1336–1340.
18. Wrobel, J. and Dietrich, A. (1993) Preparation of L-(phosphonodifluoromethyl) phenylalanine derivatives as non-hydrolyzable mimetics of O-phosphotyrosine. *Tetrahedron Lett.* **34,** 3543–3546.
19. Smyth, M. S. and Burke, T. R. Jr. (1994) Enantioselective synthesis of N-Boc and N-Fmoc protected diethyl 4-phosphono(difluoromethyl)-L-phenylalanine; agents suitable of the solid-phase synthesis of peptides containing nonhydrolyzable analogs of O-phosphotyrosine. *Tetrahedron Lett.* **35,** 551–554.

20. Qabar, M. N., Urban, J., and Kahn, M. (1997) A facile solution and solid phase synthesis of phosphotyrosine mimetic L-4-[diethylphosphono(difluoromethyl)]phenylalanine (F_2Pmp$(EtO)_2$ derivatives. *Tetrahedron* **53,** 11,171–11,178.
21. Otaka, A., Burke, T. R. Jr., Smyth, M. S., Nomizu, M., and Roller, P. P. (1993) Deprotection and cleavage methods for protected peptide resins containing 4-[(diethylphosphono)difluoromethyl]-DL-phenylalanine residues. *Tetrahedron Lett.* **34,** 7039–7042.
22. Gordeev, M. F., Patel, D. V., Barker, P. L., and Gordon, E. M. (1994) N-alpha-Fmoc-4-phosphono(difluoromethyl)-L-phenylalanine: A new O-phosphotyrosine isosteric building block suitable for direct incorporation into peptides. *Tetrahedron Lett.* **35,** 7585–7588.
23. Ottinger, E. A., Shekels, L. L., Bernlohr, D. A., and Barany, G. (1993) Synthesis of phosphotyrosine-containing peptides and their use as substrates for protein tyrosine phosphatases. *Biochem.* **32,** 4354–4361.
24. Rink, H. (1987) Solid-phase synthesis of protected peptide fragments using a trialkoxy-diphenyl-methylester resin. *Tetrahedron Lett.* **28,** 3787–3790.
25. Kaiser, E., Colescott, R. C., Bossinger, C. D., and Cook, P. I. (1970) Color test for the detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34,** 595–598.
26. Pennington, M. W. (1994) Site-specific chemical modification. In *Peptide Synthesis Protocols* (Pennington, M. W. and Dunn, B. M., eds.), Humana Press: Totowa, NJ, p. 173.

7

Expressed Protein Ligation for Protein Semisynthesis and Engineering

Zuzana Machova and Annette G. Beck-Sickinger

Summary

Over the past decade, a significant methodological development in peptide ligation strategies has been elaborated that now permits the assembly of peptides and proteins. Native chemical ligation (NCL) has been introduced to join synthetic unprotected peptides by using the chemoselective reaction between a C-terminal thioester and an N-terminal cysteine residue to result in a native peptide bond. Although this method has been applied to obtain peptides, small proteins, or protein domains (up to approx 150 residues), larger proteins could not been easily received because of the limited size of the ligated fragments. Intein technologies benefit from the opportunity to participate in the production of polypeptides with the reactive groups necessary for NCL aside from the rapid isolation of highly pure recombinant proteins. Expressed protein ligation extends the scope of NCL by overcoming the size limitation of target proteins accessible to synthesis. The intein splicing and EPL have been already proven to be useful for protein semisynthesis and for various investigations, including the studies of protein–protein interactions, segmental isotopic labeling for protein structure determination, synthesis of cytotoxic proteins, protein cyclization, and site-specific incorporation of noncanonical amino acids and biophysical probes into a protein sequence.

Key Words: Expressed protein ligation (EPL); intein; protein engineering; protein splicing; purification tag.

1. Introduction

The ability to alter protein structure and function by introducing noncanonical amino acids has great potential to enhance the understanding of proteins, to generate new tools for biomedical research, and to create novel therapeutic

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

agents. The insertion of small molecules such as fluorescent tags, spin resonance probes, or crosslinking agents at well-defined positions could lead to new insights into a large number of biochemical processes. Considerable attention is also focused on segmental isotopic labeling for protein structure determination using nuclear magnetic resonance (NMR). Throughout the last few years, several synthetic ***(1,2)*** and biosynthetic ***(3)*** approaches and strategies expanding the *E. coli* genetic code ***(4)*** have been developed for site-specific protein modification. However, substantial limitations of these techniques reduced their wide application. Native chemical ligation (NCL) is the most powerful strategy used for protein synthesis by regioselective linking of unprotected synthetic peptide blocks through a normal peptide bond ***(5)***. The coupling of synthetic peptide containing N-terminal cysteine residues and synthetic peptide C-terminal thioesters results in an amide bond at the ligation junction. Although this approach has proven to be useful for the semisynthesis of small proteins and protein domains ***(6)***, large proteins could not be achieved. The efficiency of the target protein synthesis is highly dependent on the number of peptide fragments; considerable handling losses appear with repeated purification and lyophilization steps. This review focuses on expressed protein ligation (EPL) as an extension of the NCL strategy ***(7)***. Based on protein splicing ***(8)***, this novel intein-mediated purification system permits isolation of peptide thioesters and/or peptides containing N-terminal cysteine residues recombinantly ***(9)***. Subsequent condensation of both peptide fragments proceeds under NCL conditions. The feasibility of EPL for different investigations (**Table 1**), including the studies of protein–protein interactions ***(7,10)***, segmental isotopic labeling for protein structure determination ***(11–13)***, introduction of noncanonical amino acids ***(7,14)*** and biophysical probes ***(15,16)***, and protein cyclization ***(17–19)*** are discussed extensively.

2. EPL

The use of recombinant proteins in EPL requires an efficient expression system for their production and purification, as well as the possibility of the specific generation of a C-terminal protein thioester and a protein possessing an N-terminal cysteine. Previously, fusion systems for the isolation of recombinant proteins have been engineered to remove the purification tag by treatment with proteases ***(20,21)***. Protease-mediated cleavage is not always complete because of the inaccessibility of the cleavage site. Cleavage within the target protein and additional amino acids derived from the fusion represent other drawbacks of this strategy. Moreover, the elevated temperatures required for proteolytic reaction may unfavorably affect the protein stability or activity. Recently, a novel system has been introduced for single-column and protease-

Table 1
Several Applications of EPL and Intein Splicing for Protein Semisynthesis and Engineering

Application	Model protein[a]	References
Segmental isotopic labeling	c-Abl, MBP	***(11,13,53)***
Central isotopic labeling	MBP, c-Crk-II	***(12,54)***
Generation of protein biosensor	c-Crk-II, c-Abl	***(15,58)***
Protein–protein interaction studies	c-Crk-II, c-Abl, GFP	***(15,16,58,62)***
Introduction of biophysical probes	Csk, c-Crk-II, c-Abl	***(7,15,16,28)***
Insertion of noncanonical amino acid	Csk, Src, SHP-2	***(7,14,59)***
Cyclization using TWIN system	Thioredoxin	***(19)***
Cyclization of domain in vivo	c-Crk-II, GFP	***(18,46)***
In vitro cyclization	c-Crk-II, β lactamase	***(17,50)***
Semisynthesis of cytotoxic proteins	RNase A, *Hpa*I	***(55)***
Semisynthesis and selenocysteine ligation	RNase A, azurin	***(34,66)***
Semisynthesis	T4 DNA ligase	***(30)***
Semisynthesis	Cre recombinase, GFP	***(31,60)***

[a]Abbreviations: c-Abl, Abelson protein tyrosine kinase; c-Crk-II, CT10 regulator of kinase; Csk, C-terminal Src kinase; GFP, green fluorescent protein; *Hpa*I, restriction endonuclease from *Haemophilus parainfluenzae*; MBP, maltose-binding protein; SHP-2, Src homology 2 protein tyrosine phosphatase; Src, Rous sarcoma virus kinase.

free purification of recombinant proteins that uses a self-cleavable affinity tag derived from a protein-splicing element ***(9)***.

2.1. Inteins and Protein Splicing

Protein-splicing elements, called inteins, represent the protein equivalents of the self-splicing RNA introns and catalyze their own excision from a precursor protein with the simultaneous fusion of the two flanking segments (N- and C-exteins) ***(8)*** as shown in **Fig. 1A**. Intein sequences consist of N- and C-terminal splicing regions and a homing endonuclease domain (**Fig. 1B**). Inteins can be recognized by their N-terminal cysteine or serine residue—alternatively, an N-terminal alanine residue—and by the insertion into the host protein just before cysteine, serine, or threonine. In addition, they contain characteristic conserved sequences such as the histidine-asparagine pair located at their C-terminus ***(22)***. Examination of the conserved amino acid residues at the intein-extein junctions gives some insights into the specific chemical reactions involved in protein splicing. The fact that the N-terminal residue at either splicing site is

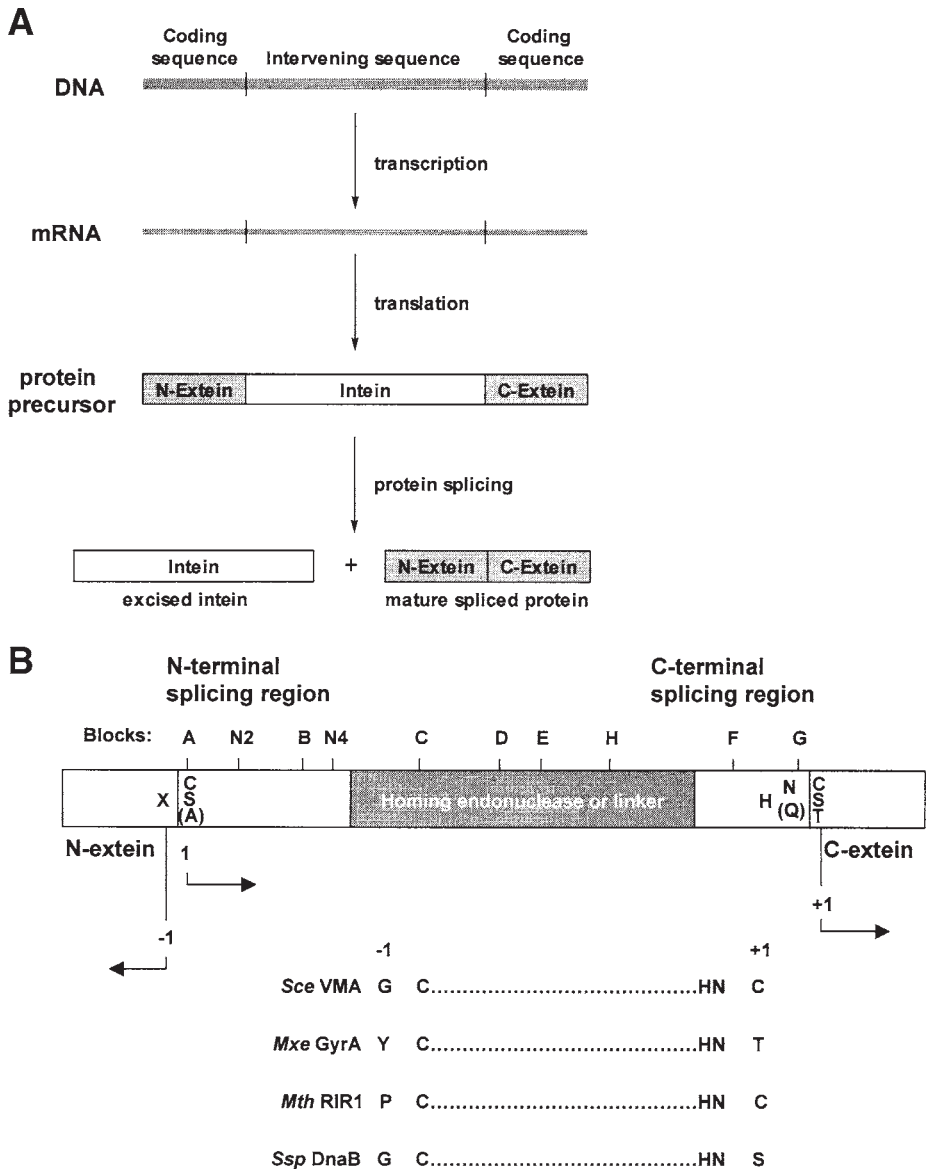


Fig. 1. Intein characteristics and mechanism of protein splicing. **(A)**, Excision of intervening sequences by protein splicing. **(B)**, Protein-splicing precursor containing an intein inserted between an N-extein (left white box) and a C-extein (right white box). Conserved sequence motifs common to all inteins (blocks A, N2, B, N4, C, D, E, H, F, and G) and conserved residues involved in splicing are depicted. Intein splicing regions (light gray) are separated by either an endonuclease domain or a short linker (dark gray). Amino acid numbering in precursor is as follows: intein amino acids are numbered from

always an amino acid with a thiol or hydroxyl side chain (cysteine, serine, threonine) suggests the involvement of ester intermediates produced by N-S or N-O acyl rearrangement in protein splicing. Indeed, individual steps of protein splicing are well-described ***(8)*** (**Fig. 2**). The identification of the residues directly participating in the breakage and formation of peptide bonds opened the way for modulating the protein splicing reaction for various applications in protein chemistry ***(23)***. Replacement of the N-terminal amino acid residue of the intein or of the Asn residue at intein's C-terminus completely blocks splicing at the corresponding termini. Since the thioester can also be achieved by transesterification in the presence of a thiol such as cysteine or dithiothreitol ***(24)***, specifically engineered mini-inteins are used for purification of protein thioester by an inducible thiol-mediated cleavage (**Table 2**).

2.2. IMPACT Purification System

Controlled cleavage at a single intein splicing junction led to development of a one-column protein purification system ***(9)***. This so-called IMPACT system, **i**ntein-**m**ediated **p**urification with an **a**ffinity **c**hitin-binding **t**ag, utilizes the inducible self-cleavage activity of an intein to separate the target protein from the affinity tag ***(25)***. The gene encoding the protein of interest is cloned into an expression vector in which the target protein is fused to the N- or C-terminus of a modified intein, and a chitin-binding domain (CBD) from *Bacillus circulans* (**Fig. 3**) ***(26)***. Following expression and affinity purification of the three-part fusion precursor, the intein is self-cleaved, releasing the protein of interest as a C-terminal thioester or N-terminal cysteine-possessing protein.

Currently, C-terminal fusion vectors containing modified mini-inteins from the genes of *Methanobacterium thermoautotrophicum* (*Mth* RIR1), *Mycobacterium xenopi gyrA* (*Mxe* GyrA), *Saccharomyces cerevisiae* (*Sce* VMA), and *Synechocystis* sp. PCC6803 (*Ssp* DnaB) are commercially available for isolation of C-terminal protein thioesters ***(25)***. All inteins carry a splicing mutation of C-terminal asparagine to alanine (**Table 2**); thus cleavage occurs only at the N-terminus of the intein. The expressed fusion protein, which is incapable of asparagine cyclization and cleavage at the C-terminal splice junction of the intein, is purified by adsorption to a chitin resin (**Fig. 3A**). The chitin-bound

---

Fig. 1. (*Continued*) the N-terminus to the C-terminus starting with 1 (conserved residues Cys, Ser, rarely Ala). N-extein amino acids are denoted as negative numbers starting with -1 (X) at the N-terminal splice junction. The amino acids in the C-extein are numbered from the N-terminus to the C-terminus, beginning with +1 (conserved residues Cys, Ser or Thr). Arrows indicate the direction of sequential numbering of residues. Conserved residues of inteins that are applied in intein-mediated purification systems are shown in the lower part of Fig. 1B.

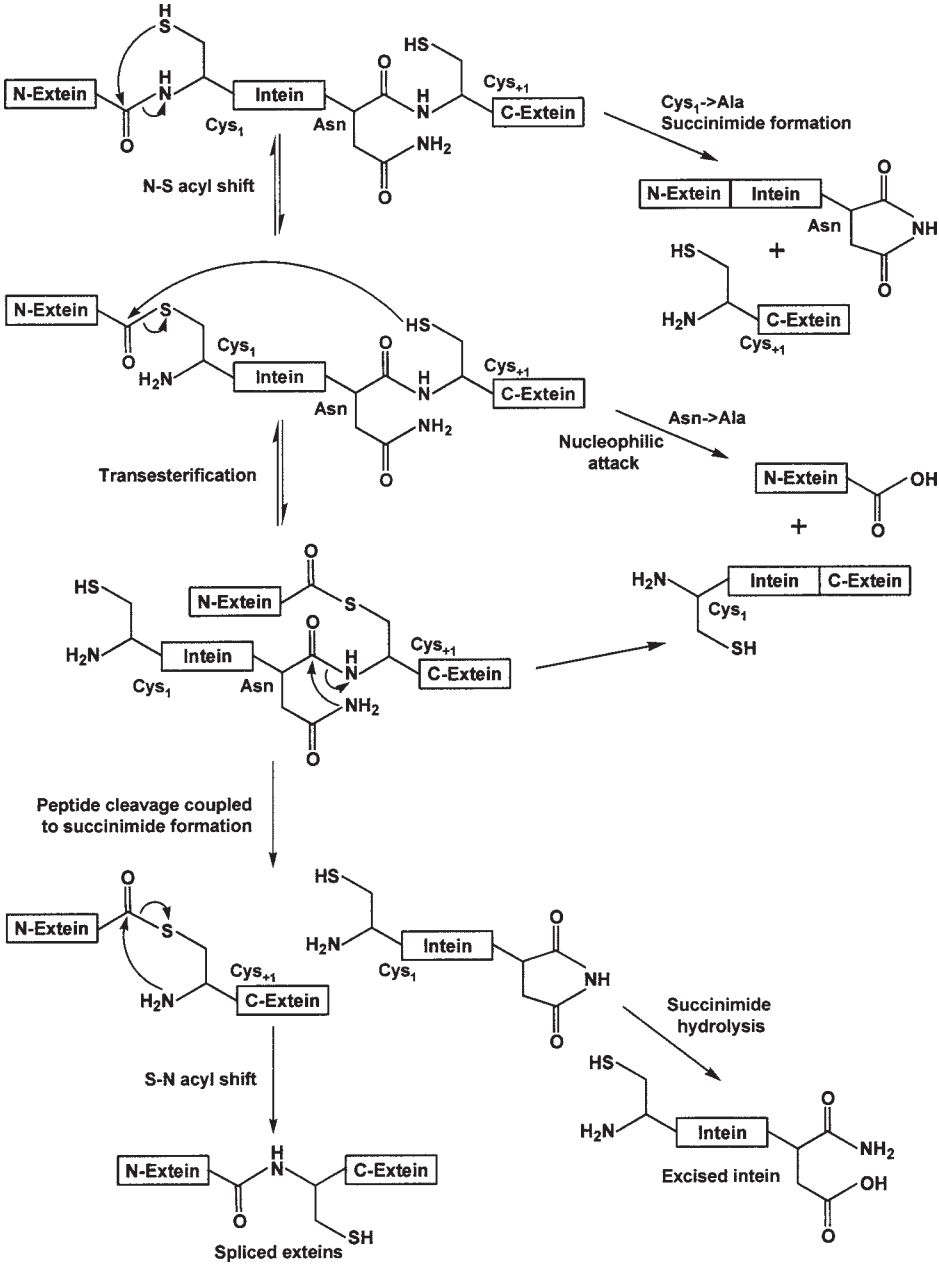


Fig. 2. Scheme of protein splicing. Cleavage pathway proposed for intein that possesses a cysteine residue in each splice junction. In the initial step a linear thioester intermediate is formed by an N-S acyl rearrangement at Cys_1 (N-terminal amino acid of the intein). Next, *trans*-thioesterification that involves nucleophilic attack of the side-

Table 2
Characterization of Intein-Mediated Purification Systems

Intein	Mutation[a]	Splice junction	Induction of cleavage	Refs.
Sce VMA	N454A	N-terminal	thiol[b]	***(9,24,37)***
Mxe Gyr A	N198A	N-terminal	thiol[b]	***(29,55)***
Mth RIR1	N134A	N-terminal	thiol[b]	***(30)***
Ssp DnaB	N154A	N-terminal	thiol[b] or pH and temperature	***(27)***
Mxe Gyr A	I2V	C-terminal	pH and temperature	***(29)***
Mth RIR1	C1A	C-terminal	pH and temperature	***(30)***
Ssp DnaB	C1A	C-terminal	pH and temperature	***(27)***

Abbreviations of inteins: *Mth* RIRI, *Methanobacterium thermoautotrophicum*; *Mxe* GyrA, *Mycobacterium xenopi* gyrase A; *Sce* VMA, *Saccharomyces cerevisiae*; *Ssp* DnaB, *Synechocystis* sp. PCC6803.

[a]Amino acids are abbreviated using single-letter notation and residues are numbered from the N-terminus of the intein.

[b]Other nucleophiles may be substituted for the thiol reagent.

precursor is unstable when treated with thiol (such as 2-mercaptoethanesulfonic acid [MESNA], dithiothreitol) or hydroxylamine, and the C-terminal thioesters are released. In the case of the *Ssp* DnaB intein the cleavage can be induced also in the absence of thiols by increasing the pH, thus allowing the purification of thiol-sensitive proteins ***(27)***.

The efficiency of the intein cleavage varies with the type of intein used for fusion and with the cleavage conditions (pH, temperature). Since usage of different thiol compounds results not only in a diverse splicing activity but also in a diverse rate of thioester hydrolysis, both parameters have to be taken into account. Specific protein properties, like the structure or the sequence at the C-terminus of the target protein, may also affect the cleavage efficiency ***(28,29)***.

Fig. 2. (*Continued*) chain of Cys_{+1} (N-terminal amino acid of the C-extein) on the thioester results in the formation of a branched intermediate. Excision of the intein occurs by peptide bond cleavage coupled to succinimide formation at the C-terminal asparagine of the intein. The ligated exteins undergo a spontaneous S-N acyl rearrangement to create a stable amide bond.

Peptide bond cleavage at either the N- or C-terminal splice junction can occur independently. Replacement of Cys_1 to Ala_1 allows only the C-terminal intein splicing and leads to C-extein that includes the cysteine at its N-terminus. Alternatively, the cleavage at N-terminus of the intein takes place by the nucleophilic attack at the thioester intermediate formed when the C-terminal Asn of the intein is substituted by Ala.

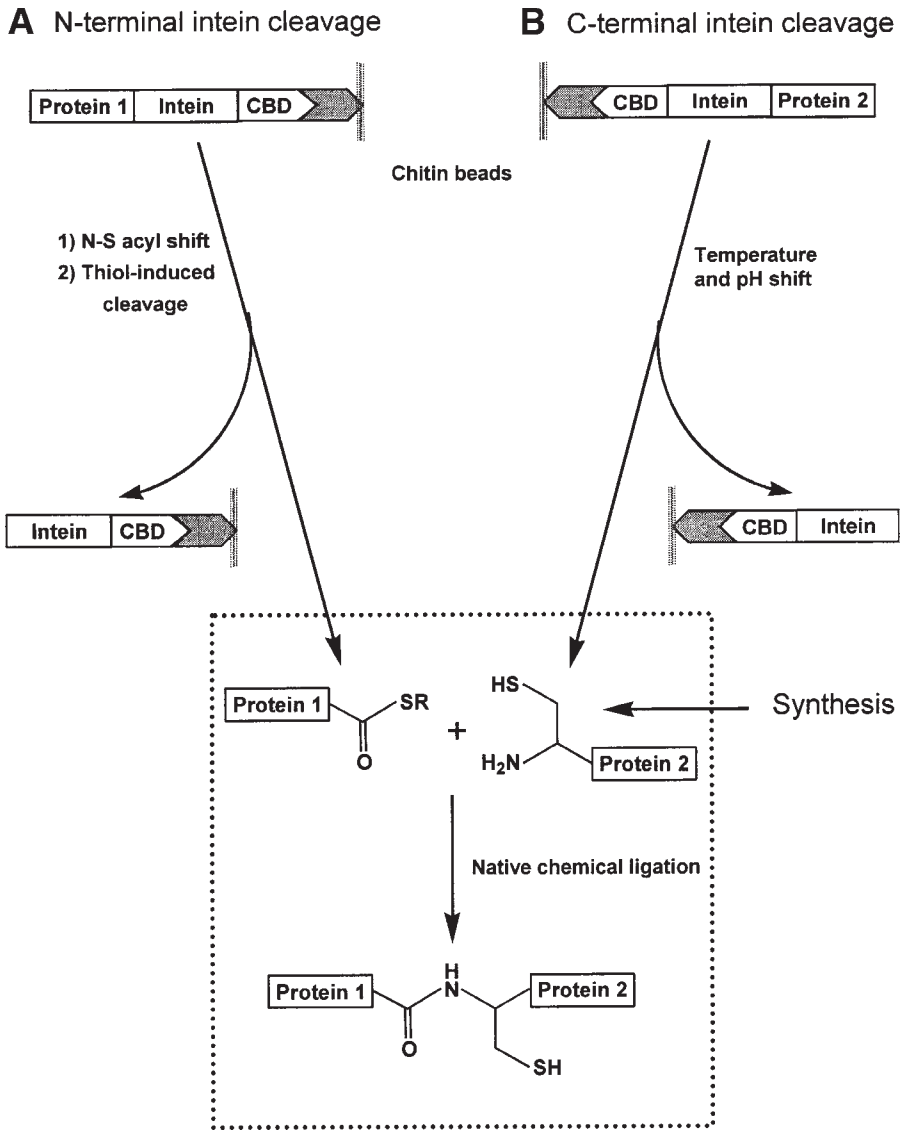


Fig. 3. Intein-mediated protein ligation. The IMPACT system allows affinity purification of proteins fused to an intein-CBD tag and their further isolation with a C-terminal thioester moiety **(A)**, or an N-terminal cysteine **(B)**. **(A)**, N-terminal intein splicing for thioester isolation. Target protein (protein 1) is expressed in *E. coli* with C-terminally located intein-CBD tag. After specific binding to the chitin resin, the thiol reagent provokes the cleavage of the peptide bond between the target protein and the intein. Whereas the intein-CBD tag remains bound to the chitin resin, the protein thioester is eluted from the column. **(B)**, C-terminal cleavage to obtain N-terminally

The modified *Mth* RIR1, *Mxe* GyrA, and *Ssp* DnaB mini-inteins have been recently applied to the isolation of proteins with an N-terminal cysteine residues ***(29,30)***. These inteins undergo temperature- and pH-dependent C-terminal cleavage when the N-terminal cysteine residue of the intein is substituted with alanine (**Table 2**). The target protein is recombinantly expressed as a fusion protein with the C-terminal intein tag ***(31)*** (**Fig. 3B**). After intein splicing the protein that possesses N-terminal cysteine is generated. Moreover, such a protein can be obtained by total chemical synthesis and different chemical labels or noncanonical amino acids can be site-specifically incorporated into the sequence.

2.3. Chemical Ligation

The principle of the reaction between the thioester and the side-chain of cysteine was reported by Wieland et al. for the first time ***(32)***. Based on these findings, Dawson et al. ***(5)*** recently introduced native chemical ligation as a remarkable improvement of the chemical ligation methods to synthesize a polypeptide with a cysteine residue at the ligation site. The chemoselective reaction of two unprotected peptides, a peptide C-terminal thioester and a peptide possessing an N-terminal cysteine, in aqueous solution at neutral pH leads to the formation of thioester-linked intermediates. Subsequent spontaneous S-N acyl rearrangement results in the native amide bond (**Fig. 4**). Alternatively, the so-called Tam's approach involves β-bromoalanine and thiocarbonic acid at acidic pH to form the covalent thioester intermediate that rapidly rearranges to give cysteine at the ligation site ***(33)***.

The ligation reaction can be performed in the presence of all other functionalities commonly found in proteins, including free cysteine sulfhydryls. Based on the NCL requirements, naturally occurring cysteine residues in the native sequence can be chosen as the site for condensation of peptide fragments. Although nonnatural cysteine residue may be inserted into the sequence, the effect of this mutation on the protein structure and function has yet to be investigated. Several other techniques, which do not require an N-terminal cysteine residue in NCL, expand the range of accessible sites of ligation to other amino acids as selenocysteine ***(34)*** or glycine ***(35)***. The incorporation of selenocysteine into protein sequences is of high interest for detailed protein structure-function

Fig. 3. (*Continued*) cysteine-possessing protein. The protein of interest (protein 2) is expressed as CBD-intein-protein 2 precursor and purified by the chitin beads. Temperature- or pH-induced intein cleavage results in proteins with an N-terminal cysteine residue. Finally, the EPL (dotted line) of the protein thioester and the cysteine containing protein, which can be also obtained synthetically, proceeds under NCL conditions.

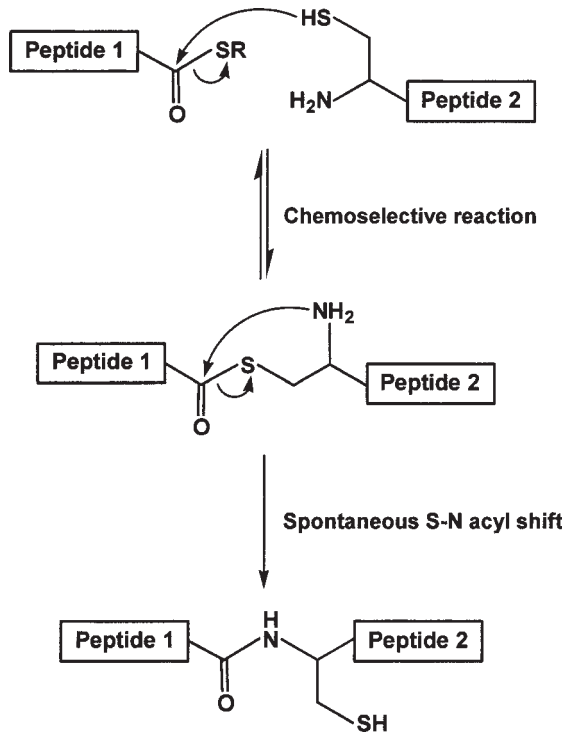


Fig. 4. Scheme of NCL. The mechanism allows the straightforward preparation of small proteins with native backbone structures from fully unprotected synthetic peptide building blocks. The initial *trans*-thioesterification step includes the chemoselective reaction between one peptide with a C-terminal α-thioester group (peptide 1) and second peptide with an N-terminal cysteine residue (peptide 2). Generated thioester-linked intermediate spontaneously rearranges to form a native peptide bond at the site of ligation.

analyses. The NCL reaction with selenocysteine is more rapid than with cysteine, especially at low pH ***(34)***. The studies with N(α)-2-phenyl ethanethiol indicated this scaffold as a suitable candidate for the development of ligation strategies that involve glycine residues at the ligation site ***(35)***. An efficient method for the generation of non-cysteine-containing proteins has been recently introduced, which makes use of alanine in the ligation ***(36)***. In order to achieve this goal, the ligation reaction is performed using the cysteine and subsequently the cysteine thiol is selectively removed by desulfurization to obtain alanine. Furthermore, conformationally assisted ligation can proceed under native conditions even without the required N-terminal cysteine ***(37)***.

Model studies indicated that all 20 naturally occurring amino acids support ligation when placed at the C-terminus of the thioester peptide; however, significant differences in the kinetics of ligation have been evaluated ***(2)***. Moreover, the nature of this amino acid influences the yield of recombinant thioester gained by using the IMPACT system ***(29,38)***.

One of the critical conditions necessary for ligation is the maintenance of a strong reducing environment to reduce disulfide formation of cysteines. Under suitable conditions, peptide thioesters undergo *trans*-thioesterification when exposed to thiol-possessing compounds. Thiol-reducing agents, such as thiophenol or MESNA, assist in the regulation of the thioester reactivity when added to the ligation mixture ***(28)***. The rate of the NCL is related to the chemical nature of the α-thioester group. Faster reactions can be obtained by using aryl thioesters ***(5)***. An important feature of all chemical ligation approaches is that they can be performed in the presence of organic solvents or chaotropic agents such as guanidine hydrochloride or urea. This often turns out to be of great practical value since many peptides and proteins are insoluble in aqueous buffers at neutral pH, particularly at the millimolar concentrations required for efficient ligation. Although extremely high concentrations of chaotrope are tolerated in NCL (up to 6 *M* urea), the efficiency of EPL dramatically drops with urea concentration higher than 4 *M* ***(68)***.

3. Protein Engineering by Expressed Protein Ligation

3.1. Site-Specific Protein Modifications

Site-directed and random mutagenesis, by which a specific amino acid of a protein can be replaced by any of the other 19 canonical amino acids, allow the generation of proteins with enhanced properties including stability, catalytic activity, and binding specificity (e.g., **ref. *39***). Nevertheless, changes in proteins are limited to the 20 canonical amino acids. In the cases where noncanonical amino acids are present in proteins, they must be generated by modifying one of the 20 amino acids after their incorporation into the protein or by altering the genetic code itself. Synthetic ***(1,2)*** and biosynthetic ***(3)*** techniques and strategies expanding the *E. coli* genetic code ***(4)*** have been applied to introduce the noncanonical amino acids into proteins. However, their broad application is limited by several factors, such as low protein yields, complex cloning, and low number of incorporated changes.

The insertion of biophysical probes into protein sequences gives an opportunity to monitor various biochemical processes, such as protein–protein interaction. In order to develop potential biosensors, disruption of the protein cooperations can be coupled to some form of signaling event, such as a fluorescence change. Ayers et al. ***(28)*** reported on the interaction between the Src homology

3 (SH3) domain of the Abelson protein tyrosine kinase (c-Abl-SH3) and its known polyproline ligand 3BP2. The insertion of two fluorophores into the ligand sequence gave the possibility to screen the protein communication by fluorescence resonance energy transfer (FRET) measurement.

Similarly, dual-labeled biosensors for phosphorylation studies of signaling protein c-Crk-II substrate by c-Abl tyrosine kinase were developed applying solid phase EPL (SPPL) strategy ***(15)***. Conformational change in c-Crk-II upon phosphorylation was monitored by FRET after sequential ligation of the target protein with the two-site, specifically labeled synthetic peptides (**Fig. 5**). Recently, a c-Crk-II biosensor with an improved characteristic was developed in the same research group ***(16)***. The significant fluorescence change upon phosphorylation permitted c-Abl kinase activity to be monitored in real-time, thus providing a tool for the screening of potential kinase inhibitors or compounds that block the interactions necessary for phosphorylation.

We focused on the semisynthesis of prohormone neuropeptide Y (proNPY) and the chemically labeled proNPY analogs ***(69)*** for further studies of prohormone processing. Two members of the prohormone convertase (PC) family, which are involved in proNPY cleavage to yield bioactive neuropeptide Y (NPY) and C-terminal peptide of NPY (CPON), could be recently identified ***(40)***. It has been suggested that the length of the substrate discriminates the activity of the processing enzymes. Based on our previous studies on the relevance of several positions within the proNPY sequence ***(41)***, five proNPY-derived analogs were synthesized; two of them contained a carboxyfluorescein label, one a biotin label, and two were without any label. Western blot analyses revealed that none of the introduced changes influenced proNPY recognition by antibodies directed against NPY or CPON, or by streptavidin.

Furthermore, we applied EPL for the semisynthesis and engineering of human interleukin 8 (hIL-8), an inflammatory CXC chemokine ***(68)***. Prospering from the occurrence of four cysteine residues (forming two intramolecular disulfide bridges) within the hIL-8 sequence, the protein was divided into two segments by using one of the cysteine residues at the ligation site. Since both disulfide bridges are important for biological activity of hIL-8 and receptor binding ***(42)***, the novel carboxyfluorescein-labeled [K^{69}(CF)]hIL-8(1-77) protein was examined in biological activity assays performed on human promyelotic HL60 cells that naturally express both hIL-8 receptor subtypes.

3.2. Backbone Cyclization and Polymerization of Proteins

Cyclization of peptides often improves their in vivo stability and biological activity ***(43)***; it is also commonly used to reduce the conformational flexibility of peptides. So far, engineering of novel disulfide bonds has been one of the most frequently applied strategies to stabilize proteins. However, the insertion

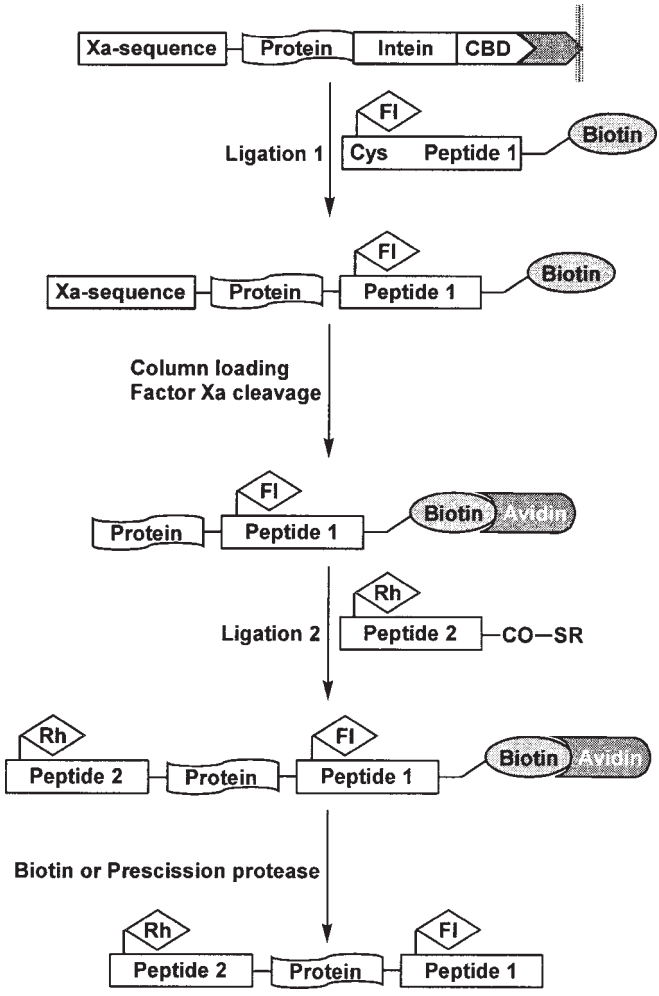


Fig. 5. Mechanism of solid phase expressed protein ligation (SPPL). Application of SPPL for the generation of a dual-labeled protein biosensor is illustrated. The expressed precursor, which includes a target protein with N-terminal cysteine (protected by a factor Xa-removable pro-sequence), is attached to the chitin matrix. Peptide 1, which contains a fluorescein probe (Fl) and a biotin affinity handle separated by a linker with recognition site for PreScission protease, is prepared synthetically. Subsequently, peptide 1 is chemoselectively ligated to the C-terminus of the target protein by using EPL. The ligation product binds to the avidin beads through its biotin functionality. In order to perform the second ligation step, the N-terminal cysteine is deprotected by factor Xa-mediated proteolysis. The newly exposed N-terminal cysteine undergoes ligation reaction with the synthetic peptide α-thioester (peptide 2) carrying a tetramethylrhodamine (Rh) probe. Finally, the dual-labeled target protein is desorbed from the solid support by biotin addition or specific cleavage with PreScission protease.

of new bonds usually interferes with the rest of the structure and complicates protein production ***(44)***. Intein-based approaches for the biosynthesis of backbone-cyclized peptides make use of the possibility to generate a circular recombinant protein by using an intramolecular version of NCL after intein splicing. In this case, the intein itself can be split and the halves, fused to the N- and C-termini of the target protein, are then reassembled ***(45–49)***. A second approach deals with EPL, when incorporation of both reactive moieties (N-terminal cysteine and an α-thioester group) proceeds within the same polypeptide and results in an efficient backbone cyclization ***(17–19,50)***.

The *trans*-splicing ability of the naturally occurring *Ssp* DnaE split intein ***(25)*** has been exploited to create a method for split intein-mediated circular ligation of peptides and proteins (SICLOPPS) ***(47)***. The expressed fusion precursor consists of the target protein inserted between C-terminal (C-intein) and N-terminal (N-intein) intein fragments (**Fig. 6A**). After spontaneous intein assembly, the standard protein-splicing reaction results in the cyclization of the target protein. The utility of this method was demonstrated in vivo and in vitro ***(45)***. Interestingly, a versatile SICLOPPS-based method for producing intracellular libraries of small cyclic peptides has been generated ***(48)***, which benefits from the possible elimination of toxic library members early in the screening process. In contrast to studies performed with naturally split intein, no linear byproduct has been detected when using artificially split intein for the purification of cyclic GFP *in vivo* ***(46)***.

The TWIN (two intein) approach has been developed for in vitro cyclization and polymerization of bacterially expressed proteins ***(19)***. Accordingly, a target protein is placed between two modified mini-inteins with either N- or C-terminal controllable cleavage activity that leads to the production of proteins with both an N-terminal cysteine and a C-terminal thioester for further EPL reaction (**Fig. 6B**). Two IMPACT-TWIN systems are commercially available, consisting of *Mxe* GyrA/*Ssp* DnaB and *Mth* RIR1/*Ssp* DnaB intein pairs. Backbone cyclization of recombinant polypeptides can be also achieved by using only one mini-intein ***(17,50)***. In the first step, the target protein is expressed with N-terminal cysteine and C-terminal intein-CBD tag. Occasionally, a leader sequence with factor Xa protease recognition site precedes the N-terminal cysteine of the protein ***(17)***. Subsequent purification on chitin beads, and possibly removal of the leader sequence, results in spontaneous ***(17)*** or thiol-mediated ***(50)*** intramolecular reaction and simultaneous cleavage of the chitin that drives to the final cyclic or polymeric product.

3.3. Segmental Isotopic Labeling

The study of biological macromolecules by nuclear magnetic resonance (NMR) spectroscopy has been greatly expanded with the use of isotopic labeling ***(51)***.

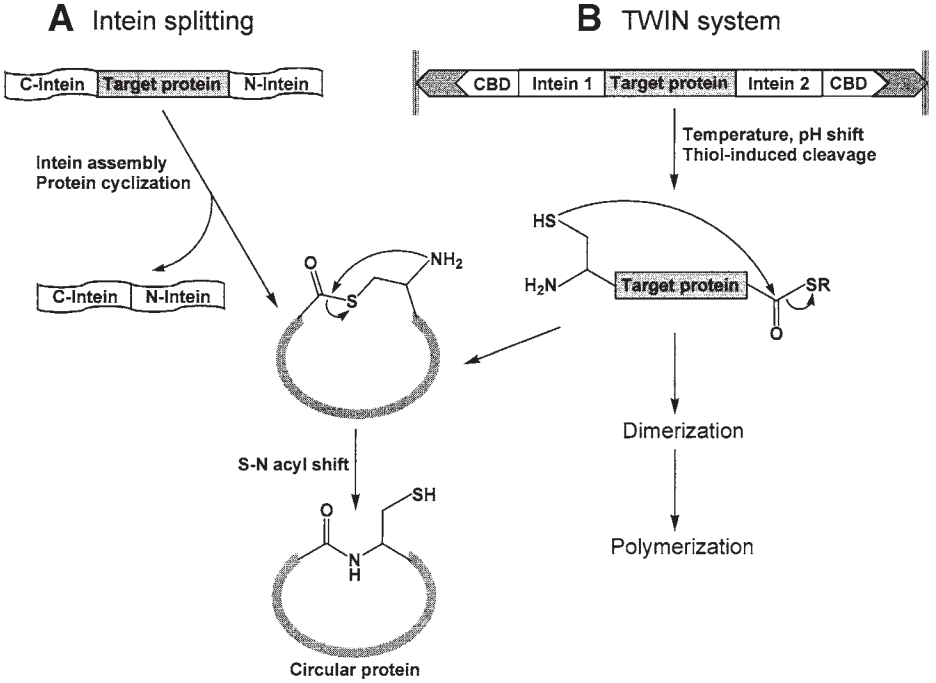


Fig. 6. Cyclization and polymerization of proteins. Two approaches that employ inteins for the generation of circular recombinant protein, split intein system **(A)**, and TWIN system **(B)**, are demonstrated. **(A)**, The target protein is inserted between the C-terminal intein (C-intein) and the N-terminal intein (N-intein) segment. After spontaneous intein assembly, the standard splicing reaction results in excised intein and cyclized target protein. **(B)**, The two intein systems sandwich the target protein between two intein-CBD tags. Controlled C- and N-terminal intein cleavages lead to target protein owning both N-terminal cysteine and C-terminal thioester. Whereas the intramolecular condensation forms cyclized proteins, intermolecular reaction gives dimeric and polymeric proteins.

Labeling of protein segments remains an important goal in general and especially in connection with the study of multidomain or modular proteins. So far, segmental isotopic labeling has been demonstrated by using two peptide ligation strategies, *trans*-splicing and EPL.

The *trans*-splicing approach is based on the reconstitution of inactive N- and C-terminal fragments of the split intein. Two recombinant protein fragments can be ligated in vitro, when each segment is expressed as a fusion protein with the complementary part of the split intein (**Fig. 7A**). The desired target protein is then generated after noncovalent association of the corresponding intein

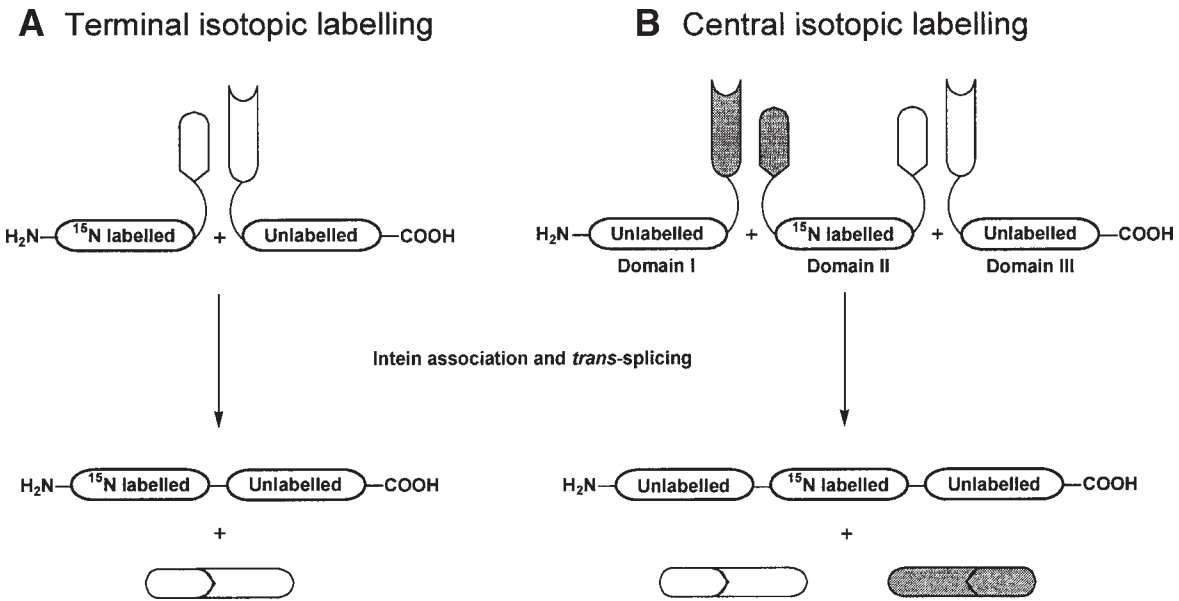


Fig. 7. Split intein approach for segmental isotopic labeling. The individual protein segments are expressed in unlabeled or isotopically enriched medium as fusion proteins carrying complementary parts of the split intein. *Trans*-splicing is achieved by reconstituting inactive N- and C-terminal intein fragments, and results in ligation of recombinant protein segments. Thus, terminally **(A)**, or centrally **(B)** labeled proteins of interest are gained. Corresponding intein fragments are illustrated in white (split intein 1) or gray (split intein 2).

fragments. In the pioneering study of Yamazaki et al. ***(52)***, individual domains of the *E. coli* RNA polymerase α subunit (αC) were selectively labeled with ^{15}N by using the PI-*Pfu*I intein from *Pyrococcus furiosus*. The comparison of NMR spectra of both labeled domains showed significant similarities with the reference spectrum of the uniformly ^{15}*N*-labeled αC protein. Various segments of ^{15}*N*- or ^{13}*C*-labeled maltose-binding protein (MBP) were also obtained by this method ***(13)***. Furthermore, the same group has presented a method for central-segment isotopic labeling, which allows the selective observation of any part of interest ***(12)***. Accordingly, the target protein was expressed as three split-intein fusions. The central protein segment contained PI-*Pfu*I and PI-*Pfu*II intein fragments at its termini, while the N- and C-terminal protein fragments carried the complementary intein parts (**Fig. 7B**). The N- and C-terminal protein fragments were expressed individually in unlabeled medium and an isotope-labeled culture was used for the production of the central segment.

Segmental isotopic labeling by using sequential EPL ***(53)*** represents an alternative approach to overcome the limitations of *trans*-splicing. Xu et al. ***(11)*** exploited the EPL strategy for single-domain labeling of c-Abl kinase. Because the structural organization and interactions between c-Abl domains are complex and difficult to elucidate, fragment labeling could support enlightening the effects of the surrounding domains on a segmentally labeled domain, or ligand binding by structure-activity-relationship (SAR) by NMR. Importantly, the sequential ligation of individual domains by EPL was also expanded into isotopic labeling of internal protein domains ***(54)***.

3.4. Production of Cytotoxic Proteins

Efficient production of target proteins in *E. coli* is often accompanied by two major problems; first, induction of protein expression can lead to cytotoxic effects, and second, the recombinant protein can be produced in suitable amounts but accumulates in inclusion bodies in the cytosol. The basis for the toxicity in many cases is unknown. It is believed to result from the overexpression of a fully active protein that competes with the cellular components and deregulates the cell physiology. Isolation of cytotoxic proteins as wild-type or mutant forms by applying EPL involves the expression of an inactive truncated form of the protein fused to the intein tag. After the ligation of the generated thioester with the synthetic peptide containing an N-terminal cysteine the amino acid sequence of target protein will be completed and the activity can be reconstituted in vitro.

Two potentially cytotoxic proteins were isolated in this manner, bovine pancreatic ribonuclease A (RNase A) and a restriction enzyme from *Haemophilus parainfluenzae* (*Hpa*I) ***(55)***. A naturally occurring cysteine residue close to the

C-terminus of the proteins was chosen as the site for fragment ligation in both enzymes. The truncated forms of these proteins displayed no detectable enzymatic activity. However, upon ligation with the synthetic peptide and further renaturation steps (in the case of RNase A), the enzymatic activity was recovered.

3.5. Studies of Protein–Protein Interactions

During the past few years, several studies were reported that utilized EPL to elucidate the role of protein phosphorylation–dephosphorylation reactions. Extensive interests are directed toward understanding the interactions that occur between different pathways and at the development of drugs that could inhibit specific protein kinases and phosphatases ***(56)***. The availability of EPL in studies of protein–protein interactions has been demonstrated by experiments performed with protein tyrosine kinase Csk. Target protein kinase can be engineered by introducing a unique non-naturally occurring amino acid into a conserved region of the enzyme's binding site. For example, Muir et al. focused on the insertion of phosphotyrosine-containing tail into Csk protein ***(7)***, which catalyzes the phosphorylation of a conserved tyrosine within the C-terminal tail of protein kinase Src. Besides the investigation of the effect of this modification on protein conformation and catalytic behavior, an incorporated C-terminal fluorescent tag served as a sensitive marker of ligation and as a probe for biochemical studies. Similarly, in the work of Cole et al. ***(57)***, a phosphotyrosine tail that carried fluorescent tag incorporated via a flexible linker was ligated to Csk. The Csk-catalyzed phosphorylation of Src was recently examined also by Wang and Cole ***(14)*** when tyrosine analogs were introduced into the Src kinase in place of the natural tail tyrosine residue. Kinase assays carried out using these Src protein substrates provided detailed insights into the mechanism of Src recognition by Csk.

The ability to insert synthetic peptides into recombinantly expressed proteins by using sequential EPL opened the possibility to develop fluorescence-based protein biosensors for the investigation of molecular processes. In principle, an appropriate fluorophore can be selectively introduced into a protein so that its fluorescence properties are dependent on the functional state of a screened procedure. Thus, a synthetic tripeptide containing the environmentally sensitive dansyl group was placed between the recombinantly derived SH3 and SH2 domains of c-Abl kinase ***(58)***. The generated protein biosensor was used to investigate the fluorescence change induced upon ligand binding to c-Abl-SH(32). Because of its potentiality to distinguish domain cooperation at low ligand concentrations, this system can effectively participate in identification of novel ligands and in characterization of protein–protein interactions that regulate c-Abl function.

Site-specific incorporation of a nonhydrolyzable phosphotyrosine analog revealed a role for phosphorylation of protein tyrosine phosphatase SHP-2 in cell signaling ***(59)***. The phosphorylated SHP-2 protein showed improved activity in catalyzing phosphate release than its nonphosphorylated counterpart.

3.6. Application of Green Fluorescent Protein in EPL

Unique properties enabled the green fluorescent protein (GFP), isolated from the jellyfish *Aequorea victoria*, to become one of the most widely studied and exploited proteins in biochemistry and cell biology (e.g., ***[60]***). In contrast to other bioluminescent molecules, the formation of the final fluorophore requires molecular oxygen and no external enzymes or co-factors ***(61)***. Enhanced GFP (EGFP) split system has been generated to detect protein–protein interactions ***(62,63)*** (**Fig. 8A**). In order to monitor protein cooperation in vivo, the N- and C-terminal halves of the *Sce* VDE ***(62)*** or *Ssp* DnaE ***(63)*** intein were fused to N- and C-terminal halves of EGFP. Each of these fusion proteins was linked to the protein of interest (protein A) and its target protein (protein B). In the case of protein–protein interaction, the closely oriented intein halves underwent correct folding and the splicing resulted in the synthesis of the mature EGFP. The extent of the protein–protein interaction was evaluated by measuring the magnitude of fluorescence intensity originated from the reconstituted EGFP. Interestingly, the detection of protein–protein interactions by using split luciferase has been lately demonstrated by the same researchers ***(64)*** (**Fig. 8B**).

In order to characterize potent carriers and to visualize cellular uptake, we have applied EPL for the fusion of amidated human calcitonin (hCT)-derived carrier peptide with EGFP ***(65)***. Although hCT and its C-terminal fragments have been shown to permeate the nasal epithelium, transport was limited to peptides up to now. EGFP thioester, which has been produced by using the IMPACT system, retained its native green fluorescence during intein splicing and EPL reaction. EGFP alone did not show any cell permeation, but ligated EGFP-[C^8] hCT8-32 conjugate revealed specific mucosal internalization. Accordingly, this system represents a promising approach of controlled drug delivery for large molecules in protein and gene therapy.

3.7. Expressed Enzymatic Ligation

The development of EPL has facilitated the production of large protein targets, but the requirement of specific N-terminal amino acids at the ligation site (cysteine ***[7]***, selenocysteine ***[66]***) reduces the general utilization of this method. Recently, we introduced a novel approach that we named expressed enzymatic ligation (EEL) for the semisynthesis of larger and chemically modified proteins that combines the advantages of the EPL with those of the substrate mimetic

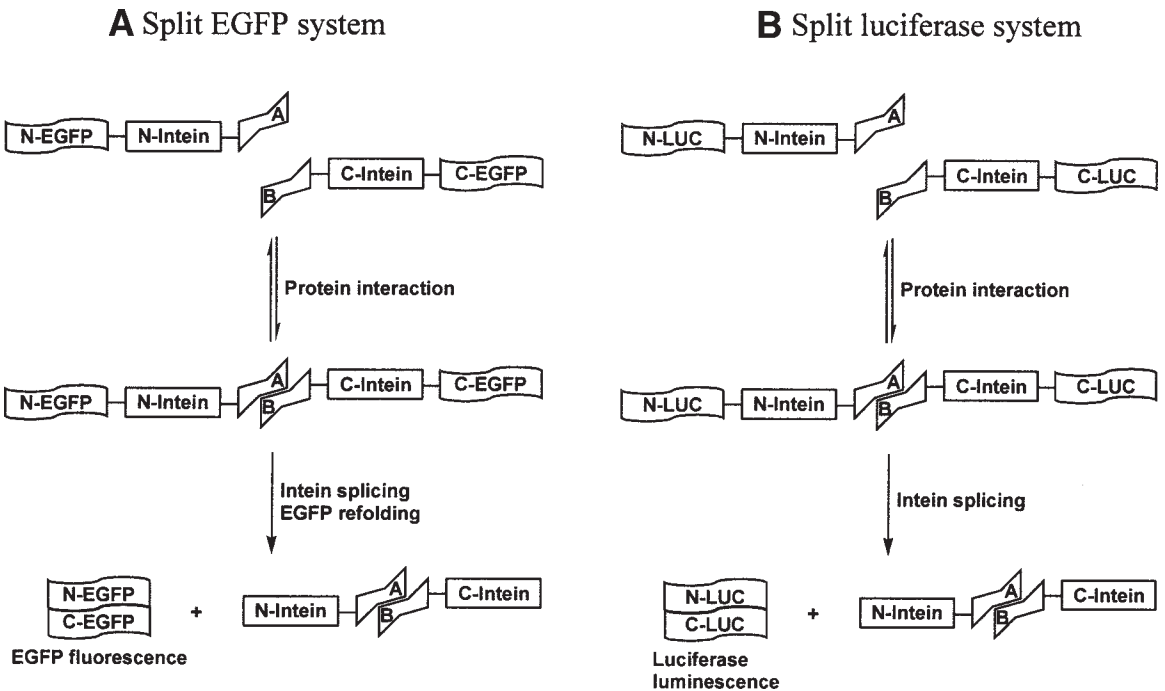


Fig. 8. Protein–protein interaction study based on split intein. In order to monitor the protein interaction in vivo, the N- and C-terminal halves of the intein (N-intein and C-intein) are fused to N- and C-terminal halves of EGFP **(A)**, or luciferase **(B)**. Each of these fusion proteins is linked to the protein of interest (protein A) and its target protein (protein B). Upon protein A–protein B cooperation, the closely oriented intein fragments mediate intein splicing. The measurement of fluorescence intensity originated from the reconstituted mature EGFP protein or measurement of luciferase luminescence is possible.

strategy ***(70)***. A commercially available protease, i.e., the Glu/Asp-specific serine protease V8 from *Staphylococcus aureus*, and simple alkyl thioesters, attainable by EPL, were used as biocatalyst and acyl donor components, respectively, in protease-catalyzed peptide ligation. Based on the concept of programming the substrate specificity of proteases ***(67)***, thioesters containing V8 protease-specific ester leaving groups were isolated by applying the IMPACT system. Subsequently, the thioesters served as substrate mimetics for the enzymatic ligation step involving several model peptides as acyl acceptor components. The ligation proceeded independently of the primary specificity of the enzyme and the nature of the acyl acceptor's N-terminal amino acid moiety. Although the V8 protease-mediated segment condensation is in an early stage of development, the great potential of the enzyme to become a useful and essential tool for protein synthesis has been demonstrated.

4. Conclusion

For many years, chemical synthesis of large peptides and proteins has been a daunting task. The expansive progression in different ligation methods assists in the successful fulfillment of this large order. Since the building blocks can be obtained efficiently either by stepwise solid-phase synthesis or by recombinant methods, the limitations of size to produce synthetic or semisynthetic proteins is no longer a concern.

Fifty years ago, Wieland et al. ***(32)*** investigated the reaction between peptide thioesters and side-chain of cysteine residues, which led to the formation of the native amide bond. Recently established native chemical ligation ***(5)*** has been proven to be a significant step for many of the latest extensions in protein semisynthesis. The method is based on the chemoselective reaction of a polypeptide containing a C-terminal thioester with a second polypeptide carrying an N-terminal cysteine residue. However, because of the restricted size of two synthetic polypeptide fragments, the applicability of this method was limited to the generation of small proteins and protein domains. Protein splicing, a process that bears remarkable similarities to NCL, offers an alternative way to obtain large protein fragments recombinantly. The intein-mediated purification system using modified inteins has become a versatile tool to accomplish the production of highly pure recombinant proteins in a single chromatographic step. Recombinant peptide thioester and N-terminally cysteine-possessing peptides are then subsequently ligated following the principle of native ligation. Thus, expressed protein ligation provides expanded opportunities in protein engineering and wide applications in areas ranging from structural biology and biochemistry to basic cell biology. Going beyond the 20 canonical amino acids, EPL permits all kind of modifications to be included in the backbone and amino acid side-chains of a protein, such as introduction of fluorophores, isotopic and

spin labels, photoactivable crosslinkers, posttranslational modifications, or various noncanonical amino acids. The feasibility of EPL has been demonstrated by a number of extensive investigations involving studies of protein–protein interactions, protein structure determination, and structure-activity studies as summarized recently ***(71)***.

References

1. Kent, S. B. (1988) Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **57,** 957–989.
2. Hackeng, T. M., Griffin, J. H., and Dawson, P. E. (1999) Protein synthesis by native chemical ligation: expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. USA* **96,** 10,068–10,073.
3. Thorson, J. S., Cornish, V. W., Barrett, J. E., Cload, S. T., Yano, T., and Schultz, P. G. (1998) A biosynthetic approach for the incorporation of unnatural amino acids into proteins. *Methods Mol. Biol.* **77,** 43–73.
4. Wang, L. and Schultz, P. G. (2002) Expanding the genetic code. *Chem. Commun.* 1–11.
5. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science* **266,** 776–779.
6. Dawson, P. E. and Kent, S. B. (2000) Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69,** 923–960.
7. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **95,** 6705–6710.
8. Paulus, H. (2000) Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* **69,** 447–496.
9. Chong, S., Mersha, F. B., Comb, D. G., et al. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192,** 271–281.
10. Severinov, K. and Muir, T. W. (1998) Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem.* **273,** 16,205–16,209.
11. Xu, R., Ayers, B., Cowburn, D., and Muir, T. W. (1999) Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc. Natl. Acad. Sci. USA* **96,** 388–393.
12. Otomo, T., Ito, N., Kyogoku, Y., and Yamazaki, T. (1999) NMR observation of selected segments in a larger protein: central-segment isotope labeling through intein-mediated ligation. *Biochemistry* **38,** 16,040–16,044.
13. Otomo, T., Teruya, K., Uegaki, K., Yamazaki, T., and Kyogoku, Y. (1999) Improved segmental isotope labeling of proteins and application to a larger protein. *J. Biomol. NMR* **14,** 105–114.
14. Wang, D. and Cole, P. A. (2001) Protein tyrosine kinase Csk-catalyzed phosphorylation of Src containing unnatural tyrosine analogues. *J. Am. Chem. Soc.* **123,** 8883–8886.

15. Cotton, G. J. and Muir, T. W. (2000) Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation. *Chem. Biol.* **7,** 253–261.
16. Hofmann, R. M., Cotton, G. J., Chang, E. J., et al. (2001) Fluorescent monitoring of kinase activity in real time: development of a robust fluorescence-based assay for Abl tyrosine kinase activity. *Bioorg. Med. Chem. Lett.* **11,** 3091–3094.
17. Camarero, J. A. and Muir, T. W. (1999) Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *J. Am. Chem. Soc.* **121,** 5597–5598.
18. Camarero, J. A., Fushman, D., Cowburn, D., and Muir, T. W. (2001) Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain. *Bioorg. Med. Chem.* **9,** 2479–2484.
19. Evans, T. C. Jr., Benner, J., and Xu, M. Q. (1999) The cyclization and polymerization of bacterially expressed proteins using modified self-splicing inteins. *J. Biol. Chem.* **274,** 18,359–18,363.
20. Riggs, P. (2000) Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol. Biotechnol.* **15,** 51–63.
21. Smith, D. B. and Johnson, K. S. (1988) Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67,** 31–40.
22. Pietrokovski, S. (1998) Modular organization of inteins and C-terminal autocatalytic domains. *Protein Sci.* **7,** 64–71.
23. Xu, M. Q. and Perler, F. B. (1996) The mechanism of protein splicing and its modulation by mutation. *Embo J.* **15,** 5146–5153.
24. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M. Q. (1996) Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J. Biol. Chem.* **271,** 22,159–22,168.
25. Xu, M. Q., Paulus, H., and Chong, S. (2000) Fusions to self-splicing inteins for protein purification. *Methods Enzymol.* **326,** 376–418.
26. Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S., and Tanaka, H. (1994) The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J. Bacteriol.* **176,** 4465–4472.
27. Mathys, S., Evans, T. C., Chute, I. C., et al. (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene* **231,** 1–13.
28. Ayers, B., Blaschke, U. K., Camarero, J. A., Cotton, G. J., Holford, M., and Muir, T. W. (1999) Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers* **51,** 343–354.
29. Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q., and Perler, F. B. (1999) Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques* **27,** 110–114, 116, 118–120.
30. Evans, T. C. Jr., Benner, J., and Xu, M. Q. (1999) The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* **274,** 3923–3926.

31. Cantor, E. J. and Chong, S. (2001) Intein-mediated rapid purification of Cre recombinase. *Protein Expr. Purif.* **22,** 135–140.
32. Wieland, T., Bokelmann, E., Bauer, L., Lang, H. U., and Lau, H. (1953) Bildung von S-haltigen Peptiden durch intramolekulare Wanderung von Aminoacylresten. *Annalen der Chemie* **583,** 129–149.
33. Tam, J. P., Lu, Y. A., Liu, C. F., and Shao, J. (1995) Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl. Acad. Sci. USA* **92,** 12,485–12,489.
34. Hondal, R. J., Nilsson, B. L., and Raines, R. T. (2001) Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **123,** 5140–5141.
35. Marinzi, C., Bark, S. J., Offer, J., and Dawson, P. E. (2001) A new scaffold for amide ligation. *Bioorg. Med. Chem.* **9,** 2323–2328.
36. Yan, L. Z. and Dawson, P. E. (2001) Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **123,** 526–533.
37. Beligere, G. S. and Dawson, P. E. (1999) Conformationally assisted protein ligation using C-terminal thioester peptides. *J. Am. Chem. Soc.* **121,** 6332–6333.
38. Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J. Biol. Chem.* **273,** 10,567–10,577.
39. Corringer, P. J., Le Novere, N., and Changeux, J. P. (2000) Nicotinic receptors at the amino acid level. *Annu. Rev. Pharmacol. Toxicol.* **40,** 431–458.
40. Brakch, N., Rist, B., Beck-Sickinger, A. G., et al. (1997) Role of prohormone convertases in pro-neuropeptide Y processing: coexpression and in vitro kinetic investigations. *Biochemistry* **36,** 16,309–16,320.
41. Pohl, R. A., Machova, Z., Söll, R., Brakch, N., Grouzmann, E., and Beck-Sickinger, A. G. (2000) Pro-NPY and truncated analogues are substrates for prohormone convertase PC1/3. *J. Pept. Sci.* **6,** S127.
42. Clark-Lewis, I., Schumacher, C., Baggiolini, M., and Moser, B. (1991) Structure-activity relationships of interleukin-8 determined using chemically synthesized analogs. Critical role of NH2-terminal residues and evidence for uncoupling of neutrophil chemotaxis, exocytosis, and receptor binding activities. *J. Biol. Chem.* **266,** 23,128–23,134.
43. Trabi, M. and Craik, D. J. (2002) Circular proteins—no end in sight. *Trends Biochem. Sci.* **27,** 132–138.
44. Li, P. and Roller, P. P. (2002) Cyclization strategies in peptide derived drug design. *Curr. Top. Med. Chem.* **2,** 325–341.
45. Evans, T. C. Jr., Martin, D., Kolly, R., et al. (2000) Protein trans-splicing and cyclization by a naturally split intein from the dnaE gene of *Synechocystis* species PCC6803. *J. Biol. Chem.* **275,** 9091–9094.
46. Iwai, H., Lingel, A., and Pluckthun, A. (2001) Cyclic green fluorescent protein produced in vivo using an artificially split PI-PfuI intein from *Pyrococcus furiosus*. *J. Biol. Chem.* **276,** 16,548–16,554.

47. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., and Benkovic, S. J. (1999) Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. USA* **96,** 13,638–13,643.
48. Scott, C. P., Abel-Santos, E., Jones, A. D., and Benkovic, S. J. (2001) Structural requirements for the biosynthesis of backbone cyclic peptide libraries. *Chem. Biol.* **8,** 801–815.
49. Siebold, C. and Erni, B. (2002) Intein-mediated cyclization of a soluble and a membrane protein in vivo: function and stability. *Biophys. Chem.* **96,** 163–171.
50. Iwai, H. and Pluckthun, A. (1999) Circular beta-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* **459,** 166–172.
51. Goto, N. K. and Kay, L. E. (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10,** 585–592.
52. Yamazaki, T., Otomo, T., Oda, N., et al. (1998) Segmental isotope labeling for protein NMR using peptide splicing. *J. Am. Chem. Soc.* **120,** 5591–5592.
53. Cowburn, D. and Muir, T. W. (2001) Segmental isotopic labeling using expressed protein ligation. *Methods Enzymol.* **339,** 41–54.
54. Blaschke, U. K., Cotton, G. J., and Muir, T. W. (2000) Synthesis of multi-domain proteins using expressed protein ligation: Strategies for segmental isotopic labeling of internal regions. *Tetrahedron* **56,** 9461–9470.
55. Evans, T. C. Jr., Benner, J., and Xu, M. Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **7,** 2256–2264.
56. Burke, T. R. Jr. and Zhang, Z. Y. (1998) Protein-tyrosine phosphatases: structure, mechanism, and inhibitor discovery. *Biopolymers* **47,** 225–241.
57. Cole, P. A., Sondhi, D., and Kim, K. (1999) Chemical approaches to the study of protein tyrosine kinases and their implications for mechanism and inhibitor design. *Pharmacol. Ther.* **82,** 219–229.
58. Cotton, G. J., Ayers, B., Xu, R., and Muir, T. W. (1999) Insertion of a synthetic peptide into a recombinant protein framework: a protein biosensor. *J. Am. Chem. Soc.* **121,** 1100–1101.
59. Lu, W., Gong, D., Bar-Sagi, D., and Cole, P. A. (2001) Site-specific incorporation of a phosphotyrosine mimetic reveals a role for tyrosine phosphorylation of SHP-2 in cell signaling. *Mol. Cell* **8,** 759–769.
60. Zhang, A., Gonzalez, S. M., Cantor, E. J., and Chong, S. (2001) Construction of a mini-intein fusion system to allow both direct monitoring of soluble protein expression and rapid purification of target proteins. *Gene* **275,** 241–252.
61. Chalfie, M. (1995) Green fluorescent protein. *Photochem. Photobiol.* **62,** 651–656.
62. Ozawa, T., Nogami, S., Sato, M., Ohya, Y., and Umezawa, Y. (2000) A fluorescent indicator for detecting protein-protein interactions in vivo based on protein splicing. *Anal. Chem.* **72,** 5151–5157.
63. Ozawa, T., Takeuchi, T. M., Kaihara, A., Sato, M., and Umezawa, Y. (2001) Protein splicing-based reconstitution of split green fluorescent protein for monitoring protein-protein interactions in bacteria: improved sensitivity and reduced screening time. *Anal. Chem.* **73,** 5866–5874.

64. Ozawa, T., Kaihara, A., Sato, M., Tachihara, K., and Umezawa, Y. (2001) Split luciferase as an optical probe for detecting protein-protein interactions in mammalian cells based on protein splicing. *Anal. Chem.* **73,** 2516–2521.
65. Machova, Z., Muhle, C., Krauss, U., et al. (2002) Cellular internalization of enhanced green fluorescent protein ligated to a human calcitonin-based carrier peptide. *ChemBioChem* **3,** 672–677.
66. Berry, S. M., Gieselman, M. D., Nilges, M. J., van Der Donk, W. A., and Lu, Y. (2002) An engineered azurin variant containing a selenocysteine copper ligand. *J. Am. Chem. Soc.* **124,** 2084–2085.
67. Cerovsky, V. and Bordusa, F. (2000) Protease-catalyzed fragment condensation via substrate mimetic strategy: a useful combination of solid-phase peptide synthesis with enzymatic methods. *J. Pept. Res.* **55,** 325–329.
68. David, R., Machova, Z., and Beck-Sickinger, A. G. (2003) Semisynthesis and application of carboxyfluorescein-labeled biologically active human interleukin-8. *Biol. Chem.* **384,** 1619–1630.
69. von Eggelkraut-Gottanka, R., Machova, Z., Grouzmann, E., Beck-Sickinger, A. G. (2003) Semisynthesis and characterisation of the first analogues of pro-neuropeptide Y. *ChemBioChem.* **4,** 425–433.
70. Machova, Z., von Eggelkraut-Gottanka, R., Wehofsky, N., Bordusa, F., and Beck-Sickinger, A. G. (2003) Expressed enzymatic ligation: a new approach for the synthesis of chemically modified proteins, Angew. *Chem. Int. Ed.* **42,** 4916–4918.
71. David, R., Richter, M. P. O., and Beck-Sickinger, A. G. (2004) Expressed protein ligation: methods and applications. *Eur. J. Biochem.* **271,** 663–677.

8

Cellular Delivery of Peptide Nucleic Acid by Cell-Penetrating Peptides

Kalle Kilk and Ülo Langel

Summary

Peptide nucleic acid (PNA) is a promising tool in biomedical research. PNA oligomers can be synthesized according to peptide synthesis protocols, but they hybridize to complementary RNA and DNA strands like oligonucleotides. The main hindrance to the use of PNA oligomers has been their poor uptake by cells. Fortunately, this limitation can be overcome by the application of peptide vectors. Different strategies are used to conjugate peptide vectors and PNA oligomers. In this chapter we discuss different types of delivery vectors (receptor ligands and cell-penetrating peptides) and strategies for the design and synthesis of PNA-vector conjugates. A protocol for the delivery of antisense PNA oligomer-cell-penetrating peptide conjugates is also described.

Key Words: Peptide nucleic acid; delivery; cell-penetrating peptides; transferrin; antisense.

1. Introduction

Peptide nucleic acid (PNA) oligomers are DNA/RNA analogs that can be synthesized according to peptide synthesis protocols. PNA synthesis was made possible by the replacement of the phosphorylated sugar moiety of nucleotides with aminoethylglycine linkers ***(1,2)*** (**Fig. 1**). The possibility of combining PNA with peptides, as well as several special properties arising from the unnatural backbone, has made PNA a very interesting tool for biomolecular studies. The lack of negative charge and flexibility of the backbone allow PNA to bind complementary DNA or RNA more tightly than conventional oligonucleotides ***(3, 4)***. Hence, PNA oligomers are able to sterically block a specific region in mRNA

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

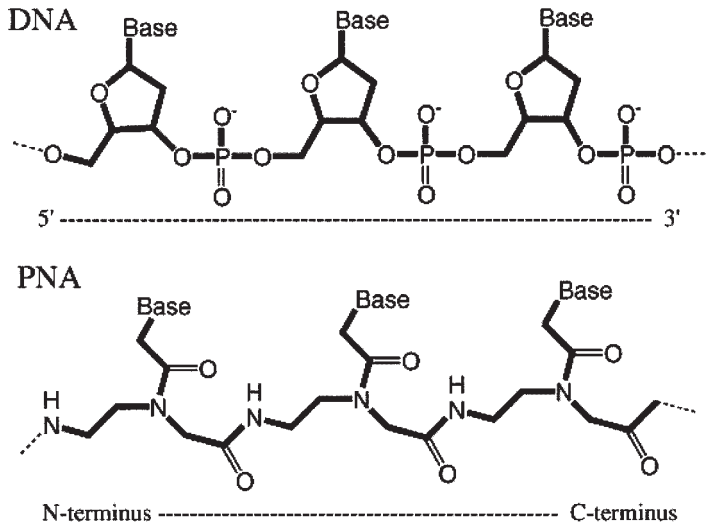


Fig. 1. Comparison of DNA and PNA structures. Base indicates any of the four nucleobases.

and thereby downregulate its translation (antisense strategy). Moreover, PNA oligomers can invade double-stranded DNA ***(5)*** to interfere with gene function or attach a label to a very specific site on the gene ***(6,7)***. Recently, PNA-transport vector conjugate-based strategies for gene delivery have been developed by several research groups ***(8,9)***. Excellent reviews are available concerning different aspects of PNA delivery and their applications ***(10–12)***.

The major obstacle to PNA use so far has been poor uptake by most cells and tissues. Therefore, delivery systems based mainly on the use of peptide vectors have been developed. Recent studies, on the other hand, demonstrate that neuronal tissues or cells are an exception to this general rule and spontaneously take up PNA ***(13–16)***. The reason for this is not clear, but the fact is that if injected directly into neuronal tissue, or applied to a neuronal cell culture, a delivery vector is not obligatory for cellular uptake of PNA oligomers.

Peptide vectors for PNA can be divided into two major categories: (1) cell surface receptor ligands and (2) cell-penetrating peptides. Receptor-mediated uptake is based on the fact that some receptors (e.g., transferrin ***[9,17]***) are subject to endocytosis following agonist activation. PNA, or any other cargo coupled to the agonist, will thereby also be internalized via endocytosis. Cell-penetrating peptides (CPPs) have been known since 1994, when penetratin was discovered (*see* **ref. *18***). All CPPs possess the ability to penetrate the cell membrane by a seemingly receptor- or energy-independent mechanism, and most if not

all of them can carry cargoes. The exact mechanism of CPP entry is still unknown, but CPPs have proved to be efficient delivery vectors for oligonucleotides ***(19)***, proteins ***(20)***, and even liposomes ***(21)***. CPPs and their applications are extensively reviewed in a recent book ***(22)***, and more information is also provided in Chapter 5 of this volume. To date, transportan ***(23,24)***, penetratin ***(14,23,25)***, and Tat ***(25)*** have been used for PNA delivery.

One may ask about the differences between CPPs and receptor ligands as delivery vectors for PNA and other cargoes. In this respect, several important issues are noted here. The first issue to consider is that of cell targeting. Only cells expressing the appropriate receptor, allowing targeted delivery, take up a PNA-ligand conjugate. Conversely, the CPPs developed so far are not capable of distinguishing between cell types. Second, the delivery vector should be as inert as possible and not interfere with the cargo's effect under investigation. Activation of a receptor will certainly trigger a cascade of intracellular signals that could complicate any interpretation of results. Some CPPs may also induce unwanted intracellular effects. pVEC is an example of a CPP for which no considerable intracellular activity has yet been assigned ***(26)***.

Additional factors that could influence the delivery efficiency and possible side effects of PNA-vector chimeras include the physical linkage between the PNA sequence and vector. Below we provide an overview of different conjugation strategies and provide a step-by-step protocol for disulfide bridge formation between PNA and carrier peptide.

Lipofection is a possible alternative to the more common peptide-mediated delivery of PNA. However, the neutral hydrophilic backbone of PNA makes it unlikely to associate with lipofection reagents. The use of semicomplementary DNA oligomers that bind to both PNA and, via charged–backbone interactions, to lipofection reagents may circumvent this problem. The drawback to this strategy is that for each PNA sequence a separate DNA oligomer must be synthesized. Moreover, the melting temperature between PNA and DNA must conform to certain limits that, according to Braasch and Corey ***(27)***, are within the range of 50 to 70°C. Complexes with lower melting temperatures disassemble before translocating across the cell membrane and very stable complexes inactivate the PNA. In this book, which is devoted to peptides, we do not provide a detailed protocol for lipofection, though such information is available elsewhere ***(27)***.

2. Materials

2.1. PNA Synthesis

The following reagents are from Applied Biosystems (Foster City, CA) and should be stored at –20°C.

1. *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU).
2. PNA monomers: *t*-Boc-A(Z)-OH; *t*-Boc-C(Z)-OH; *t*-Boc-G(Z)-OH; *t*-Boc-T-OH. Or: Fmoc-A(Bhoc)-OH; Fmoc-C(Bhoc)-OH; Fmoc-G(Bhoc)-OH; Fmoc-T-OH.
3. Equipment and materials for common peptide synthesis (*see* Chapter 1).

2.2. Conjugation of PNA and Peptides Via a Disulfide Bridge

1. Cys containing PNA and peptide, one of which must have 3-nitro-2-pyridinesulphenyl (Npys)-derivatized Cys.
2. 200 µL dimethylsulfoxide (DMSO).
3. 200 µL dimethylformamide (DMF).
4. 100 µL acetate buffer 0.1 M, pH 5.5.
5. Solvents for reverse phase HPLC purification:
   a. Eluent A: 99.9% acetonitrile (AcN) + 0.1% trifluoroacetic acid (TFA).
   b. Eluent B: 99.9% H_2O + 0.1% TFA.
   c. C_{18} column semipreperative scale HPLC column.

2.3. Cellular Delivery of CPP-PNA Conjugates for Antisense Applications

1. Cultured cells at about 60–80% confluence in growth medium.
2. Culture medium appropriate for selected cell line.
3. 1 m*M* Peptide-PNA conjugate solution in water.
4. Equipment for Western blotting or alternative method to determine protein concentration.

3. Methods

3.1. Design of PNA-Peptide Conjugates

An important question is how to achieve the most efficient delivery of PNA. This problem does not usually have a definite answer and many different factors must be considered. **Table 1** reviews the main types of PNA-peptide conjugates and indicates the most important issues in each case. The internalization routes are schematically drawn in **Fig. 2**. Depending on the peptide attached, the PNA oligomer is internalized via endocytosis or directly delivered across the cell membrane. For nuclear localization signals (NLS) and tetralysine (K_4) the delivery route is not clear yet. The same holds true for nuclear delivery with CPPs. Some CPPs (e.g., transportan) have been shown to accumulate in the nucleus, while others have not.

The designer must not forget strategically important sites in the peptide. For example, it is believed that the free N-terminal is important for the internalization of transportan and, therefore, the side-chain of Lys in position 13 is used for cargo attachment ***(28)***. It is believed, although not clearly demonstrated

Table 1
Possible Methods to Improve PNA Delivery Via Attachment of Peptides, With Comments on Applicability

Structure	Comments: Pros	Comments: Cons	Ref.
PNA-$(Lys)_{1-4}$	✓ Simplest way to improve PNA delivery	✓ Efficiency not fully validated yet	***(30)***
PNA-CPP[a]	✓ Does not distinguish between cells ✓ Some CPPs do not have intracellular side effects	✓ Possible synthesis complications in longer (> 20 building blocks) sequences	***(31)***
PNA-ligand	✓ Uptake depends on receptor expression allowing selective targeting	✓ Possible synthesis complications in longer (>20 building blocks) sequences ✓ Receptor is activated/ blocked—affects cell signaling	
PNA-S-S-peptide[b]	✓ Easier to synthesize in case of long peptides and PNAs ✓ Disulfide bridge is degraded in cytoplasm—peptide does not interfere with PNA or its target	✓ Needs additional Cys residues in PNA and peptide ✓ Additional step—disulfide bridge formation required	***(9,23,32)***
PNA-NLS[c]	✓ Improved nuclear delivery	✓ Improvement in cellular delivery is not fully validated	***(8)***
PNA-NLS-S-S-peptide	✓ Cytoplasmic delivery and improved nuclear delivery	✓ Most complicated synthesis	***(14)***

[a]Here and later: whether PNA is attached to N- or C-terminus of peptide depends on structural requirements for active peptide. Wrongly attached PNA can abolish peptide interactions with membrane or receptor.

[b]Here and later: "peptide" means either CPP or receptor ligand.

[c]NLS is a nuclear localization signal, usually a highly positively charged heptamer: PKKKRKV from SV40 core protein.

yet, that the separation of the carrier and the PNA oligomer after internalization is important for the functionality of PNA. For the synthesis of biodegradable conjugates, [$N^{13\varepsilon}$-Cys(Npys)]-transportan or [$N^{1\alpha}$-Cys(Npys)]-pVec are the most recommended vector peptides. As methods for disulfide bridge formation

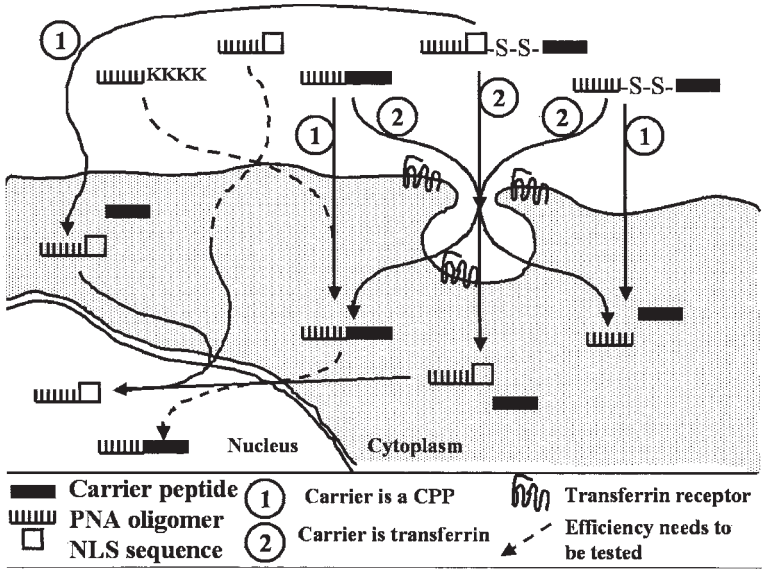


Fig. 2. Possible ways to deliver a PNA oligomer over cell- and nuclear membrane by peptides. Transferrin receptor mediated uptake (2) is endocytotic; CPP-mediated uptake (1) goes via a seemingly energy-independent route. The efficiency of some steps (shown in dashed line) is not completely proven and may be very low under certain experimental conditions.

and cellular treatment with the aforementioned vectors are similar, we do not specify the carrier peptide in the protocols below.

3.2. PNA Synthesis

Like amino acids, PNA monomers have a carboxyl group and an amino group. Thus, they can be assembled into peptide chains like ordinary amino acids. Monomers with *t*-Boc or Fmoc protection on the amine are commercially available from, e.g., Applied Biosystems, and the different considerations for PNA synthesis are well described ***(29)***. Steps and reagents not commonly used in peptide synthesis but important for PNA synthesis are as follows:

1. Start the synthesis by downloading the capacity of the resin in order to avoid interactions between growing PNA chains that can terminate the synthesis (*see* **Note 1**). The solid support should have substitution of between 0.01 and 0.2 mmol/g. The majority of commercial resins have higher substitutions than this, making downloading necessary.
2. Swell commercial resin in DCM for at least 30 min and wash with DMF, DCM, DMF, DCM, DCM, DCM.

3. Preactivate 0.01–0.2 mmol of the first amino acid/PNA monomer per each gram of resin. For PNA, HATU (0.9 eq to monomer) is the most recommended activator, but TBTU can also be used (*see* **Note 2**).
4. Couple the first building block to the resin in 50% DMF/DCM in the presence of a molar equivalent of the proton acceptor diisopropylethylamine (DIEA). Coupling times vary from 30 min for most of the amino acids to 60 min for Val, Ile, and Arg and PNA monomers.
5. Wash: EtOH, DMF, 5% DIEA in DCM, DMF, DCM, DCM, DCM.
6. Block any free amino groups with Ac_2O/NMP/Pyr 1/25/25 in the presence of DIEA in 5 min.
7. Wash: DMF, 5% DIEA in DCM, DMF, 5% DIEA in DCM, DMF, DCM, DCM, DCM.
8. Verify coupling efficiency with the Kaiser test (*see* Chapter 1). Repeat the acetylation step if beads are blue.

Actual downloading of the resin can be measured by a quantitative ninhydrine test. Normally it yields 90–100% of theoretical value.

The synthesis continues like common solid phase peptide synthesis, including final deprotection and cleavage.

Protective groups on PNA side chains are compatible with HF and TFA/H_2O cleavage in *t*-Boc- and Fmoc- chemistries, respectively. A few aspects to consider are as follows:

- In manual synthesis, use at least 30 min coupling time for PNA monomers.
- Prefer HATU to other coupling reagents.
- In *t*-Boc chemistry, the use of 5% *m*-cresol in TFA for deprotection increases synthesis yield. After capping with acetic anhydride/pyridine/*N*-methylpyrrolidone (1:25:25), wash with 10% piperidine in DMF to fully remove the capping reagent (optional step).

3.3. Conjugation of PNA and Peptide Via a Disulfide Bridge

In order to obtain a heterodimeric disulfide bridge the cysteine residue of one component, either PNA or peptide, must be derivatized. 3-Nitro-2-pyridinesulphenyl (NPys)-derivatized Cys is specifically reactive toward free thiols. Npys-labeled Cys is commercially available and with special cautions (*see* **Note 3**) can be assembled into a peptide chain like a commonly protected amino acid.

1. Weigh 1 molar eq (0.5–2 mg) of peptide and 1 molar eq of PNA in separate microcentrifuge tubes. The coupling efficiency varies between sequences and depends on relative solubilities and purities; therefore a 1:1 molar ratio may not be optimal in every case.
2. Dissolve PNA in 200 µL deoxygenized DMSO (*see* **Note 4**).
3. Dissolve peptide in 100 µL of 0.01 *M* acetate buffer, pH 5.5.
4. Add 200 µL of DMF to either of the solutions (*see* **Note 4**).

5. Mix the two solutions and vortex thoroughly.
6. Stir the mixture overnight, or at least for 4 h, at room temperature. Protect from light.
7. Separate reaction products by semipreparative RP-HPLC.
   a. Use a C_{18} column.
   b. Gradient: isocratic 20% eluent A for 5 min, followed by a linear increase of eluent A to 100% in 40 min. 20% acetonitrile at the start does not let unreacted PNA interact with stationary phase in the column and it is eluted together with organic solvents (DMSO and DMF). Conjugated PNA precedes the peptide peak (*see* **Note 5**).
   c. Detection wavelengths are 218 nm, absorbance maximum for peptide bonds, and 260 nm, which detects PNA nucleobases. (For a single wavelength detector use 260 nm.)
   d. Collect fractions absorbing at both wavelengths.
   e. Freeze dry fraction(s) and store in the dark at −20°C. Mass spectrometric analysis of conjugate is an optional step to further verify the desired product. Care must be taken not to reduce the disulfide bridge in preparation of sample or collection of mass spectra.

3.4. Cellular Delivery of CPP-PNA Conjugates for Antisense Applications

1. 24 h before experiments, passage cells so that the next day they reach 40 to 60% confluence.
2. On the day of the experiment, replace the culture medium with fresh medium (*see* **Note 6**) (optional step).
3. Prepare a fresh 1 m*M* solution of PNA-peptide conjugate or thaw an aliquot of stock solution. Repeated freeze-thaw cycles should be avoided.
   a. A serial dilution from 1 m*M* can be carried out to evaluate lower concentrations.
   b. 10 µL of PNA-peptide conjugate solution is used per mL of media (per 1 well in a 24-well plate). Whether the turnover rate of the target protein is known or not, it is good practice to screen different time points. five time points each in duplicate requires 100 µL of conjugate solution.
4. Add 10 µL of 1 m*M* solution per 1 mL of media upon cell culture. Final concentration will be 10 µ*M* of peptide-PNA conjugate. This is the highest concentration that we can recommend because of possible side effects and/or toxicity of the conjugate.
5. After appropriate time points (e.g., 1 h, 6 h, 12 h, 24 h, and 36 h), harvest the cells and quantify protein levels with an appropriate assay (e.g., Western blot). In the case of a protein with known half-life, the time schedule might be different. Nevertheless, the correlation between maximum effect of PNA and protein half-life is not well defined as yet; thus, a time scan is always informative (*see* **Note 7**).

4. Notes

1. A lysine in the C-terminus of PNA improves PNA solubility and is more suitable for resin downloading than PNA monomers.

2. The *t*-Boc-G(Z)-OH monomer takes time to dissolve in DMF or NMP; addition of DMSO aids dissolution. Slight warming, but avoiding temperatures over 50°C, can also be useful.
3. Since the Npys group is sensitive to thiols, Npys-containing peptides cannot be treated with thiophenol (used for deprotection of His(DNP) side chain) or cleaved with *p*-thiocresol, ethanedithiol, and other thiol containing scavengers.
4. Organic solvents in disulfide bridge formation protocols increase PNA solubility. If the PNA solubility in buffer is known to be high, their amount can be reduced. If PNA is insoluble in organic solvents, 20 µL of TFA may be added to the pellet.
5. If the peptide has similar hydrophobicity to the PNA and eluates at AcN concentrations <40%, another gradient or eluent system might be required for efficient separation.
6. Uptake is considerably better in the absence of serum compared with full growth medium. Serum-free medium alone must be replaced with serum-containing medium 4 h after PNA treatment.
7. If PNA effects are measured a long time after the first treatment, additional quantities of PNA-peptide construct can be exogenously added to maintain intracellular PNA concentration. This is the case when the target mRNA of the PNA sequence encodes a protein with a long half-life.

References

1. Nielsen, P. E., Egholm, M., Berg, R. H., and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254(5037),** 1497–1500.
2. Nielsen, P. E., Egholm, M., and Buchardt, O. (1994) Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone. *Bioconjug. Chem.* **5(1),** 3–7.
3. Egholm, M., Buchardt, O., Christensen, L., et al. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **365(6446),** 566–568.
4. Giesen, U., Kleider, W., Berding, C., Geiger, A., Orum, H., and Nielsen, P. E. (1998) A formula for thermal stability (Tm) prediction of PNA/DNA duplexes. *Nucleic Acids Res.* **26(21),** 5004–5006.
5. Peffer, N. J., Hanvey, J. C., Bisi, J. E., et al. (1993) Strand-invasion of duplex DNA by peptide nucleic acid oligomers. *Proc. Natl. Acad. Sci. USA* **90(22),** 10,648–10,652.
6. Zelphati, O., Liang, X., Nguyen, C., et al. (2000) PNA-dependent gene chemistry: stable coupling of peptides and oligonucleotides to plasmid DNA. *Biotechniques* **28(2),** 304–310, 312–314, 316.
7. Broude, N. E., Demidov, V. V., Kuhn, H., et al. (1999) PNA openers as a tool for direct quantification of specific targets in duplex DNA. *J. Biomol. Struct. Dyn.* **17 (2),** 237–244.
8. Branden, L. J., Mohamed, A. J., and Smith, C. I. (1999) A peptide nucleic acid-nuclear localization signal fusion that mediates nuclear transport of DNA. *Nat. Biotechnol.* **17(8),** 784–787.

9. Liang, K. W., Hoffman, E. P., and Huang, L. (2000) Targeted delivery of plasmid DNA to myogenic cells via transferrin-conjugated peptide nucleic acid. *Mol. Ther.* **1(3),** 236–243.
10. Koppelhus, U. and Nielsen, P. E. (2003) Cellular delivery of peptide nucleic acid (PNA). *Adv. Drug Deliv. Rev.* **55(2),** 267–280.
11. Nielsen, P. E. (1999) Applications of peptide nucleic acids. *Curr. Opin. Biotechnol.* **10(1),** 71–75.
12. Braasch, D. A. and Corey, D. R. (2002) Novel antisense and peptide nucleic acid strategies for controlling gene expression. *Biochemistry* **41(14),** 4503–4510.
13. Tyler, B. M., McCormick, D. J., Hoshall, C. V. et al. (1998) Specific gene blockade shows that peptide nucleic acids readily enter neuronal cells in vivo. *FEBS Lett.* **421(3),** 280–284.
14. Braun, K., Peschke, P., Pipkorn, R., et al. (2002) A biological transporter for the delivery of peptide nucleic acids (PNAs) to the nuclear compartment of living cells. *J. Mol. Biol.* **318(2),** 237–243.
15. Adlerz, L., Soomets, U., Holmlund, L., Viirlaid, S., Langel, Ü., and Iverfeldt, K. (2003) Down-regulation of amyloid precursor protein by peptide nucleic acid oligomer in cultured rat primary neurons and astrocytes. *Neurosci. Lett.* **336(1),** 55–59.
16. Rezaei, K., Xu, I. S., Wu, W. P., et al. (2001) Intrathecal administration of PNA targeting galanin receptor reduces galanin-mediated inhibitory effect in the rat spinal cord. *Neuroreport* **12(2),** 317–320.
17. Pardridge, W. M., Boado, R. J., and Kang,Y. S. (1995) Vector-mediated delivery of a polyamide ("peptide") nucleic acid analogue through the blood-brain barrier in vivo. *Proc. Natl. Acad. Sci. USA* **92(12),** 5592–5596.
18. Lindgren, M., Hällbrink, M., Prochiantz, A., and Langel, Ü. (2000) Cell-penetrating peptides. *Trends Pharmacol. Sci.* **21(3),** 99–103.
19. Oehlke, J., Birth, P., Klauschenz, E., et al. (2002) Cellular uptake of antisense oligonucleotides after complexing or conjugation with cell-penetrating model peptides. *Eur. J. Biochem.* **269(16),** 4025–4032.
20. Pooga, M., Kut, C., Kihlmark, M., et al (2001) Cellular translocation of proteins by transportan. *FASEB J.* **15(8),** 1451–1453.
21. Tseng, Y. L., Liu, J. J., and Hong, R. L. (2002) Translocation of liposomes into cancer cells by cell-penetrating peptides penetratin and tat: a kinetic and efficacy study. *Mol. Pharmacol.* **62(4),** 864–872.
22. Langel, Ü., ed. (2002) *Cell-Penetrating Peptides: Processes and Applications.* CRC, Boca Raton, FL.
23. Pooga, M., Soomets, U., Hällbrink, M., et al. (1998) Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat. Biotechnol.* **16(9),** 857–861.
24. Kaushik, N., Basu, A., Palumbo, P., Myers, R. L., and Pandey, V. N. (2002) Anti-TAR polyamide nucleotide analog conjugated with a membrane-permeating peptide inhibits human immunodeficiency virus type 1 production. *J. Virol.* **76(8),** 3881–3891.

25. Koppelhus, U., Awasthi, S. K., Zachar, V., Holst, H. U., Ebbesen, P., and Nielsen, P. E. (2002) Cell-dependent differential cellular uptake of PNA, peptides, and PNA-peptide conjugates. *Antisense Nucleic Acid Drug Dev.* **12(2),** 51–63.
26. Elmquist, A., Lindgren, M., Bartfai, T., and Langel, Ü. (2001) VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions. *Exp. Cell Res.* **269(2),** 237–244.
27. Braasch, D. A. and Corey, D. R. (2001) Synthesis, analysis, purification, and intracellular delivery of peptide nucleic acids. *Methods* **23(2),** 97–107.
28. Pooga, M., Hällbrink, M., Zorko, M., and Langel, Ü. (1998) Cell penetration by transportan. *FASEB J.* **12(1),** 67–77.
29. Nielsen, P. E. and Egholm, M. (ed.) (1999) *Peptide Nucleic Acids: Protocols and Applications.* Horizon Scientific, Wymondham, UK.
30. Sazani, P., Gemignani, F., Kang, S. H., et al. (2002) Systemically delivered antisense oligomers upregulate gene expression in mouse tissues. *Nat. Biotechnol.* **20(12),** 1228–1233.
31. Aldrian-Herrada, G., Desarmenien, M. G., Orcel, H., et al. (1998) A peptide nucleic acid (PNA) is more rapidly internalized in cultured neurons when coupled to a retro-inverso delivery peptide. The antisense activity depresses the target mRNA and protein in magnocellular oxytocin neurons. *Nucleic Acids Res.* **26(21),** 4910–4916.
32. Villa, R., Folini, M., Lualdi, S., Veronese, S., Daidone, M. G., and Zaffaroni, N. (2000) Inhibition of telomerase activity by a cell-penetrating peptide nucleic acid construct in human melanoma cells. *FEBS Lett.* **473(2),** 241–228.

9

Quenched Fluorescent Substrate-Based Peptidase Assays

Rebecca A. Lew, Nathalie Tochon-Danguy, Catherine A. Hamilton, Karen M. Stewart, Marie-Isabel Aguilar, and A. Ian Smith

Summary

The use of specific quenched fluorescent substrates (QFS) provides a rapid and sensitive method to measure peptidase activity, and is readily adaptable to high-throughput screening of potential peptidase inhibitors. In this chapter, we discuss general considerations for the development of QFS assays, and describe in detail an assay protocol for the mammalian metallopeptidase, endothelin-converting enzyme.

Key Words: Quenched fluorescence substrate; peptidase; inhibitor; endothelin-converting enzyme; bradykinin.

1. Introduction

Cleavage of bioactive peptides by specific peptidases, either to inactive products or to fragments with altered activities, represents a critical point in the regulation of peptide levels. As such, peptidases, together with proteases, represent a major target group for pharmaceuticals. Full characterization of peptidases ideally includes detailed information on their distribution, cellular and subcellular localization, and substrate specificity, as well as the design of specific inhibitors, for either therapeutic and/or research purposes. Thus, the development of sensitive and specific activity assays, particularly those that may be adapted for a high-throughput format, are highly desirable. One of the most useful strategies is that of using quenched fluorescent substrates (QFS), in which a short peptide sequence that is recognized and cleaved by the peptidase of interest (preferably with some degree of specificity) is flanked by a matched fluorophore and quencher pair. Such a pairing takes advantage of fluorescent resonance energy transfer (FRET), in which light emitted from the

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

fluorophore is effectively absorbed by the quencher moiety. This intramolecular quenching is efficient over short distances, generally less than 50 Å; upon cleavage of the intervening peptide sequence, the fragment containing the quencher diffuses away, resulting in an increase in fluorescence emanating from the fluorophore. Quenched fluorescent substrates are particularly useful for endopeptidases or proteases that recognize sequences C-terminal to the cleavage site (S1', S2', etc), and thus cannot be assayed using simpler substrates with C-terminal fluorescent leaving groups.

Assays using quenched fluorescent substrates have many advantages over alternative techniques used to monitor peptidase activity, such as HPLC. These include relative speed and ease of assays, the ability to continuously monitor fluorescence during the course of an enzyme reaction, and high sensitivity, all of which facilitate the development of high-throughput assays. However, a number of caveats exist, and each QFS assay protocol must be rigorously verified before use. Depending on the intended purpose of the QFS, its specificity for the peptidase of interest may be crucial. For example, if one wishes to measure activity of one peptidase in crude biological samples (e.g., tissue extracts, plasma, urine), then specificity is absolutely necessary to avoid nonspecific fluorescence following substrate cleavage by other enzymes. Specificity can usually be achieved by careful design of the QFS, keeping in mind that hydrolysis at any of the peptide bonds will result in fluorescence. Should absolute specificity be impossible, it may be feasible to minimize the action from other enzymes, for example, by performing the assay at a pH that allows only the enzyme of interest to be active, or by including specific inhibitors of other enzymes in the assay. In some cases, specificity may not be necessary; for example, screening of potential inhibitors against a recombinant or purified enzyme. Other potential difficulties include solubility problems, high background fluorescence (arising either from contaminants in the QFS preparation or inefficient quenching), interference from sample proteins, and intermolecular quenching at high substrate concentrations. For the most part, all these problems can be overcome or compensated for if identified in the development of the assay.

In this chapter, we will describe our attempts to design a specific QFS for the endothelin-converting enzyme, ECE-1. This peptidase catalyzes the final step in the production of a potent 21-amino-acid vasoconstrictor, endothelin (ET), from a 38-residue precursor, bigET. Specific inhibitors of ECE-1 are of potential therapeutic use, as excess ET production is associated with hypertension and cardiac hypertrophy. However, the design of inhibitors has been hampered by the close homology of ECE-1 with another membrane-bound peptidase, neutral endopeptidase (NEP). The development of a specific ET-based QFS for ECE-1 has also proven to be a challenge, in that the enzyme requires an extended sequence of bigET for efficient cleavage. Included in this chapter are

details of the design and analysis of ECE-1 substrates, based on both the bigET sequence and another peptide substrate, bradykinin (BK), as well as a step-by-step protocol for our current ECE QFS assay.

2. Materials

1. Quenched fluorescent substrates (QFS). The QFS listed in **Tables 1** and **2** were synthesized by standard solid phase peptide synthesis techniques, and purified by HPLC (*see* **Notes 1** and **2**). The identity of each QFS was verified by mass spectrometry. The BK-based QFS (compound no. 26, **Table 2**), published previously by others ***(1)***, was solubilized in DMSO:ethanol (50:50) at a stock concentration of 2 mg/mL (=1.2 m*M*), and kept at −20°C (*see* **Note 3**). Stock QFS was diluted 1:10 in assay buffer just prior to use.
2. Assay buffer: 50 m*M* Tris-HCl, 150 m*M* NaCl, pH 6.3 (*see* **Note 2**).
3. Enzyme source (*see* **Note 4**): Chinese hamster ovary cells (CHO-K1, American Type Culture Collection) were transiently transfected via standard DEAE-dextran methods with an expression construct encoding the human ECE-1c protein. The medium was changed to serum-free medium the day after transfection. Two days after transfection, cells were harvested and a crude membrane fraction sedimented by ultracentrifugation (100,000*g* for 1 h). The pellet was suspended in assay buffer, aliquoted, and frozen at −70°C until assay. Protein concentration was determined by standard methods (e.g., Lowry or BCA). The equivalent of 3 µg total protein was used in each assay well.
4. Microtiter plates: black, round-bottomed 96-well plates (Nunc).
5. Fluorimeter: *f*Max from Molecular Devices (Sunnyvale, CA), SoftMax analysis software (*see* **Note 5**).

3. Methods

1. Turn on fluorimeter, set thermostat at 37°C, and set fluorescence filters at 320 nm excitation and 420 nm emission (*see* **Note 6**).
2. Add assay buffer to wells at appropriate volume (*see* **Table 3**). All samples, including blanks, are performed in triplicate.
3. Add any inhibitors to wells (**Table 3**)—generally, 12 µL of a 10X concentration.
4. Add enzyme preparation (in this example, 8 µL ECE-1c membranes).
5. Pre-incubate enzyme and inhibitors at 37°C for 30 min in *f*Max.
6. Add diluted QFS (40 µL) to all wells (**Table 3**) (*see* **Note 3**).
7. Immediately take a time zero reading (*see* **Note 7**).
8. Incubate in *f*Max.
9. Read fluorescence at designated time intervals (*see* **Note 8**).
10. Perform data analysis (*see* **Note 9**).

4. Notes

1. In the design of a QFS to assay peptidase activity, both the peptide sequence and the fluorophore/quencher pair must be chosen carefully. Ideally, the peptide sequence

Table 1
Substrates for Endothelin-Converting Enzyme Based on Natural Substrate Big Endothelin (BigET)

Compound	Name	Sequence	% Cleavage	Fluorescent yield
		Nonfluorescent substrates		
1	BigET(18–34)	DIIW^VNTPEHVVPYGLG-NH_2	60	
2	[F^{22}]-BigET(18–34)	DIIW^**F**NTPEHVVPYGLG-NH_2	70	
3	BigET(19–34)	_IIW^VNTPEHVVPYGLG-NH_2	60	
4	BigET(20–34)	__IW^VNTPEHVVPYGLG-NH_2	<5	
5	BigET(18–33)	DIIW^VNTPEHVVPYGL_-NH_2	5	
6	[$F^{21,22}$]-BigET(18–34)	DII**F**^**F**NTPEHVVPYGLG-NH_2	<5	
7	[$F^{21,22}$]-BigET(19–34)	_II**F**^**F**NTPEHVVPYGLG-NH_2	70	
		Quenched Fluorescent substrates		
8	MCA-[K(DNP)^{35}]-BigET(18–34)	**MCA-**DIIW^VNTPEHVVPYGLG**K(DNP)**-NH_2		Low
9	K(DNP)-[F^{22}, K(MCA)^{35}]-BigET(18–34)	**K(DNP)**-DIIW^**F**NTPEHVVPYGLG**K(MCA)**-NH_2		Low
10	E(Edans)-[F^{22}, K(Dabcyl)^{35}]-BigET(19–34)	**E(Edans)-**IIW^**F**NTPEHVVPYGLG**K(Dabcyl)**-NH_2		Low
11	D(Edans)-[F^{22}, K(Dabcyl)^{35}]-BigET(19–34)	**D(Edans)-**IIW^**F**NTPEHVVPYGLG**K(Dabcyl)**-NH_2		Low
12	D(Edans)-[F^{22}, K(Dabcyl)^{35}]-BigET(18–34)	**D(Edans)-**DIIW^**F**NTPEHVVPYGLG**K(Dabcyl)**-NH_2		Low
13	[Y^{21}, F(pNO_2)^{22}]-BigET(19–34)	_II**Y**^**F(pNO_2)**NTPEHVVPYGLG-NH_2	>95	Medium
14	[F(pNO_2)^{22}]-BigET(19–34)	_IIW ^**F(pNO_2)**NTPEHVVPYGLG-NH_2	70–80	Medium
15	[F(pNO_2)^{22}]-BigET(18–34)	DIIW^ **F(pNO_2)**NTPEHVVPYGLG-NH_2	70–80	Medium

Single letter code is used to represent amino acid residues; MCA = 7-methoxycoumarin–4-acetyl; DNP = 2, 4-dinitrophenyl. Site of cleavage is indicated by ^.

Table 2
Substrates for Endothelin-Converting Enzyme Based on Natural Substrate Bradykinin (BK)

Compound	Name	Sequence	% Cleavage
16	Bradykinin (BK)	RPPGF^SPFR	60
17	[Ala^1]-BK	**A**PPGF^SPFR	40
18	[Ala^2]-BK	R**A**PGF^SPFR	40
19	[Ala^3]-BK	RP**A**GF^SPFR	40
20	[Ala^4]-BK	RPP**A**F^SPFR	40
21	[Ala^5]-BK	RPPG**A**^SPFR	40
22	[Ala^6]-BK	RPPGF^**A**PFR	70
23	[Ala^7]-BK	RPPGF^S**A**FR	>95
24	[Ala^8]-BK	RPPGF^SP**A**R	<5
25	[Ala^9]-BK	RPPGF^SPF**A**	70
26	MCA-[Ala^7, K(DNP)^9]-BK Fluorescent substrate—high fluorescent yield	**MCA-**RPPGF^SAF**K(DNP)**	70

Single letter code is used to represent amino acid residues; MCA = 7-methoxycoumarin–4-acetyl; DNP = 2, 4-dinitrophenyl. Site of cleavage is indicated by ^.

Table 3
Assay Reagent Volumes for Endothelin-Converting Enzyme Assay

Well	Buffer	Enzyme	Inhibitor	QFS
Blank	80 µL	–	–	40 µL
Enzyme control	72 µL	8 µL	–	40 µL
Enzyme + inhibitor	60 µL	8 µL	12 µL	40 µL

should be as short as possible so that the distance between the fluorophore and quencher is minimized. Short sequences also reduce the cost of synthesis and often aid in the solubility of the compound. However, in the case of ECE-1 substrates based on bigET, we found that the minimal sequence necessary for efficient cleavage (as assessed by reverse-phase HPLC) was 16–17 residues long (Compounds 1–3, **Table 1**), similar to that seen by others ***(2–4)***. Attachment of a standard fluorophore/quencher pair, MCA and DNP, to the N- and C-termini of sequence no. 1 resulted in a substrate that was well-cleaved, but the increase in fluorescence was very low (no. 8). We hypothesized that the tryptophan N-terminal to the cleavage site was partially quenching the MCA; however, replacement of this residue with Phe (not shown), or exchanging the N- and C-terminal positions of the fluorophore and quencher (no. 9), did not increase the fluorescence generated. Indeed, similar compounds incorporating a different fluorophore-quencher pair, Edans/

Dabcyl, also showed a marked basal fluorescence in the absence of cleavage (no. 10–12); thus quenching is inefficient at these distances. Attempts to reduce the distance between quencher and fluorophore by introducing a *p*-nitrophenylalanyl residue (Phe(pNO$_2$)) into the sequence and using either the native Trp^{21} or a substituted Tyr^{21} as a fluorophore improved the fluorescence yield somewhat (no. 13–15), but an assay using these substrates would still be too insensitive for our purposes. A QFS based on another ECE-1 peptide substrate, BK, was published by Johnson and Ahn ***(1)***; our data verify their finding that substitution of Pro^7 with Ala^7 increased the rate of cleavage of bradykinin by ECE-1 (no. 23, **Table 2**), and that the addition of an N-terminal MCA and a C-terminal Lys(DNP) resulted in a useful QFS for screening potential inhibitors of ECE (no. 26). However, this substrate is not specific for the enzyme—it can be cleaved by a number of other zinc metallopeptidases, including neutral endopeptidase, angiotensin-converting enzyme, thimet oligopeptidase, and neurolysin. This lack of specificity precludes using the BK-based QFS to measure ECE-1 catalytic activity in tissue or plasma samples, but is currently being used to screen potential inhibtors of recombinant ECE-1.

2. As with any assay, conditions for a QFS assay must be optimized in terms of buffer composition, particularly pH and potential salt and metal ion effects. Furthermore, the specificity of cleavage should be verified by an independent method, such as reverse-phase HPLC. Given the aromatic nature of both fluorophore and quencher, even low levels of QFS and cleavage products can be easily detected by UV absorbance. This method will also determine whether any absence of fluorescence is due to lack of cleavage or to a low fluorescent yield, as occurred with some of the ECE substrates discussed in **Note 1**.
3. Some QFS compounds are difficult to solubilize at high concentration in aqueous solution. Dissolving in DMSO or ethanol can circumvent this, but care must be taken as these solvents may interfere with enzyme activity. For example, we have determined that final concentrations of DMSO or ethanol above 2% significantly inhibit endothelin converting enzyme activity.
4. The source of the peptidase will depend on the aim of the assay; i.e., one may wish to measure catalytic activity in tissue or plasma samples, in which case specificity of the substrate is of prime importance. Conversely, if the assay is to be used in the development of peptidase inhibitors, the enzyme should be present in as pure a form as possible. We routinely express a recombinant form of ECE-1 in Chinese hamster ovary cells, and generally use the wild-type, membrane-bound form of the enzyme in crude membrane preparations. However, we have also designed a soluble, secreted form of ECE-1 containing a hexahistidine tag; this allows purification on a nickel affinity resin and eliminates potential problems involved with the use of crude, particulate membranes.
5. Useful features in a fluorescence detector include a 96-well plate format, internal thermostatted chamber, variable wavelengths for excitation and emission, and user-friendly software for analysis of kinetic assays. For some applications, an older-style, single-cuvet fluorimeter can be used, but this is completely impractical when

assaying large numbers of samples, and requires a greater amount of both sample and QFS (as well as patience). Likewise, although variable wavelength capability provides flexibility, it is not a necessary feature, provided the filters available are appropriate for the specific fluorophore used (*see* **Note 6**).

6. Excitation and emission wavelengths will depend on the fluorophore used, and should be optimized as much as is practical. The excitation and emission spectra of most common fluorophores are readily available (e.g., from Molecular Probes); if a variable-wavelength fluorimeter or fluorescence plate reader is used, then optimization is easy. If only fixed wavelengths are available, select those closest to the optima and verify that the fluorescence generated is sufficient for the intended assay.
7. Some quenched fluorescence substrates may emit significant levels of background fluorescence in the absence of proteolytic cleavage. This may be caused by inefficient quenching of intact substrate (as occurs with increasing peptide chain length) or contaminating fluorescent byproducts of the synthesis. Provided this background fluorescence is low, blank wells, containing only QFS in buffer, should be included in the assay, and simply subtracted from the fluorescence generated in the presence of enzyme. Should background fluorescence increase with time, buffers should be replaced, as they are likely to be contaminated by microbial proteases.
8. One great advantage of quenched fluorescence substrates is the ability to perform continuous assays, as no derivitization steps are necessary (as with some fluorogenic or chromogenic substrates). However, this is only practical using a thermostatted plate reader, such as the *f*Max, rather than individual cuvet-style fluorimeters. Determining fluorescence at multiple time points not only generates more information about the enzymatic reaction, but also means that preliminary optimization experiments are not necessary for samples with unknown activity.
9. QFS-based assays lend themselves readily to standard kinetic analysis of enzyme reactions, owing to their high sensitivity (allowing detection of very low levels of product formation), the ability to continuously monitor substrate cleavage over time (*see* **Note 8**), and the easy adaptation to a high-throughput, 96-well format. Using the fluorescent product (e.g., MCA-RRPGF in the case of the bradykinin-based ECE substrate), or even free fluorophore (provided there are no structural elements of the true product which might alter fluorescence) to generate a standard curve allows the rapid determination of Michaelis-Menten parameters (K_m and V_{max}) of purified enzymes, as well as inhibition constants (K_i) of potential protease inhibitors. As kinetic analysis involves measurement of cleavage velocity over a range of substrate concentrations, it is important to ascertain whether intermolecular quenching occurs at high QFS concentrations. This is most readily done by measuring standard product fluorescence in the presence of increasing concentrations of intact substrate. Ideally, kinetic data should be compared to separate standard curves of product at each substrate concentration; however, this can be costly in terms of substrate used. If interference at high substrate concentrations is minimal, then analysis can proceed using a single product standard curve. However, if quenching of product fluorescence is significant, then alternate means of analysis, such as HPLC (*see* **Note 2**), should be pursued.

References

1. Johnson, G. D. and Ahn, K. (2000) Development of an internally quenched fluorescent substrate selective for endothelin-converting enzyme-1. *Anal. Biochem.* **286,** 112–118.
2. Okada, K., Arai, Y., Hata, M., Matsuyama, K., and Yano, M. (1993) Big endothelin-1 structure important for specific processing by endothelin-converting enzyme of bovine endothelial cells. *Eur. J. Biochem.* **218,** 493–498.
3. Claing, A., Neugebauer, W., Yano, M., Rae, G. A., and D'Orleans-Juste, P. (1995) [Phe22]-big endothelin-1[19–37]: a new and potent inhibitor of the endothelin-converting enzyme. *J. Cardiovasc. Pharmacol.* **26(Suppl 3),** S72–S74.
4. Keller, P. M., Lee, C. P., Fenwick, A. E., Atkinson, S. T., Elliott, J. D., and DeWolf, W. E. Jr. (1996) Endothelin-converting enzyme: substrate specificity and inhibition by novel analogs of phosphoramidon. *Biochem. Biophys. Res. Commun*. **223,** 372–378.

10

A Convenient Method for Synthesis of Cyclic Peptide Libraries

Gregory T. Bourne, Jonathon L. Nielson, Justin F. Coughlan, Paul Darwen, Marc R. Campitelli, Douglas A. Horton, Andreas Rhümann, Stephen G. Love, Tran T. Tran, and Mark L. Smythe

Summary

Cyclic peptides have been reported to bind to multiple, unrelated classes of receptor with high affinity. Owing to the robustness of amide bond chemistry, the ability to explore extensive chemical diversity by incorporation of unnatural and natural amino acids, and the ability to explore conformational diversity, through the incorporation of various constraints, arrays of cyclic peptides can be tailored to broadly sample chemical diversity. We describe the combination of a safety catch linker with a directed-sorted procedure for the synthesis of large arrays of diverse cyclic peptides for high-throughput screening.

Key Words: Cyclicpeptides; combinatorial synthesis; safety-catch linker.

1. Introduction

Small cyclic peptides represent a large class of biologically relevant molecules ***(1–14)***. Their interest to medicinal chemists rests with their ability to mimic biologically relevant regions of protein diversity, such as β-turns and loop motifs ***(2)***. Compared to their linear precursors they are more stable to degradative peptidases, more bioavailable, and possess entropic advantages within molecular recognition. Some cyclic peptides are drugs in their own right. For example, cyclosporin A is a cyclo-undecapeptide that is clinically used as an immunosuppressant in the treatment of autoimmune disorders and for preventing organ transplant rejection ***(3)***, while tyrocidine A and gramicidin S are both cyclo-

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

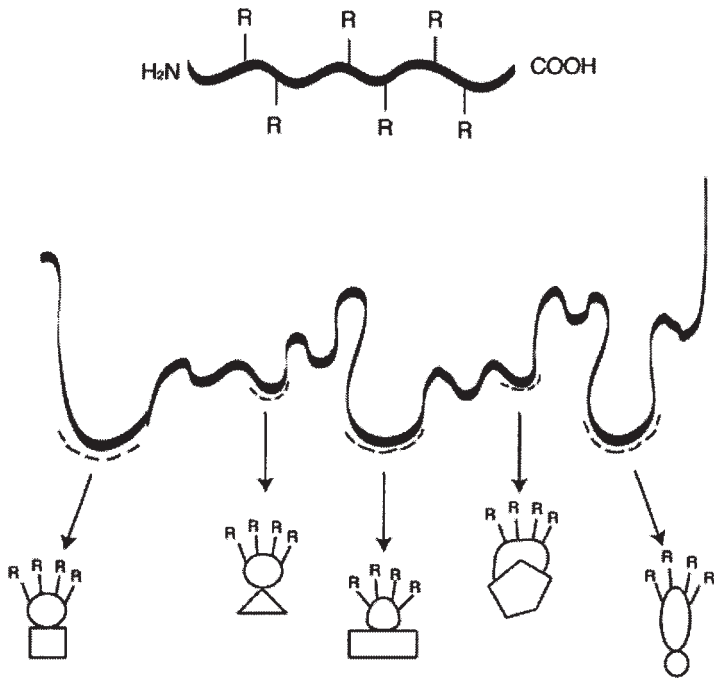


Fig. 1. Schematic illustration of a series of constrained cyclic peptides collectively sampling the broad conformational space of a linear analog. The rectangles, triangles, etc. represent organic scaffolds that force peptides into different conformational substates.

decapeptide antibiotics (***4***,**5**). Cyclic peptides have also been found to be ion channel adapters ***(6)***, as somatostatin agonists (***7,8***), neurokinin-1 and -2 receptor antagonists ***(7,9)***, δ-opioid receptor antagonists ***(10)***, $\alpha_4\beta_1$ integrin receptor antagonists ***(11)***, histone deacetylase inhibitors ***(12,13)*** and as antifungals that act through the inhibition of glucan synthesis (the candins) ***(14)***, to name a few.

1.1. Importance of Cyclic Peptide Libraries

Libraries of cyclic peptides may be exceedingly useful as "molecular toolkits" for identifying leads in the drug discovery process. Constraining a single set of pharmacophoric groups into different conformational substates (as illustrated in **Fig. 1**) provides a valuable library for probing various receptors. Such a library is "optimally diverse," as it explores both the conformational and the chemical elements of diversity. Active and inactive analogs from such a library allow one to deduce a pharmacophore for biological activity, which could be subsequently used to drive a small-molecule development program.

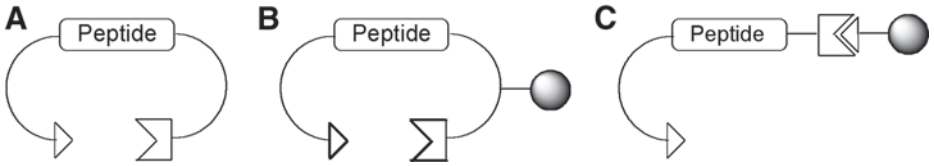


Fig. 2. Methods for the cyclization of head-to-tail cyclic peptides. Method A, solution cyclization. Method B, cyclization via side chain or backbone attachment. Method C, cyclization via cleavage through an active ester.

1.2. Strategies for Cyclic Peptide Synthesis

The appeal of cyclic peptides as a rich source of biologically active molecules makes the synthesis of combinatorial libraries of these compounds desirable. However, few avenues for the synthesis of large arrays of cyclic peptides exist. This is primarily caused by the difficult orthogonal deprotection requirements, which require a careful choice of synthetic strategy. For example, if a solution-phase head-to-tail cyclization is undertaken ***(15)*** (**Fig. 2A**), the peptide must be purified at each step of the synthesis (i.e., after synthesis of the linear, cyclized protected and after deprotection) (*see* **Note 1**).

It is therefore not surprising that solid phase synthesis of cyclic peptides is the most desirable route, and several strategies have been used for this purpose (**Fig. 2B,C**). One of the most common methods for synthesizing cyclic peptides on the solid support is by attachment of the peptide through a side chain functionality ***(16–25)*** (**Fig. 2B**). This strategy requires orthogonal protection at the C-terminus (usually accomplished by an allyl, fluorenylmethyl, or a benzyl protecting group) and requires a lysine, histidine, cysteine, tyrosine, threonine, serine, aspartic or glutamic acid, for example, in the peptide sequence to link the peptide to the solid support. Cyclization is achieved on resin, and side-chain deprotection and cleavage occur in one step.

A minor disadvantage of this approach is that it necessitates the presence of an amino acid with appropriate side-chain functionality in the sequence. This can be circumvented through the use of backbone amide linkers **1**, **2**, and **3** through which any backbone amide nitrogen (except for proline) may be attached to resin (**Fig. 3**) ***(26–28)***. In addition to this, backbone amide substitution is known to inhibit solid-phase aggregation and can aid in the cyclization of difficult sequences through promotion of the *cis*-amide bond conformation (resulting from reduction of the *cis-trans* amide bond rotational energy barrier).

Another common strategy involves activated linkers, which allow cyclization/cleavage directly from resin (**Fig. 2C**). A selection of activated linkers **4**, **5**, and **6** that have previously been used for combinatorial synthesis is shown in

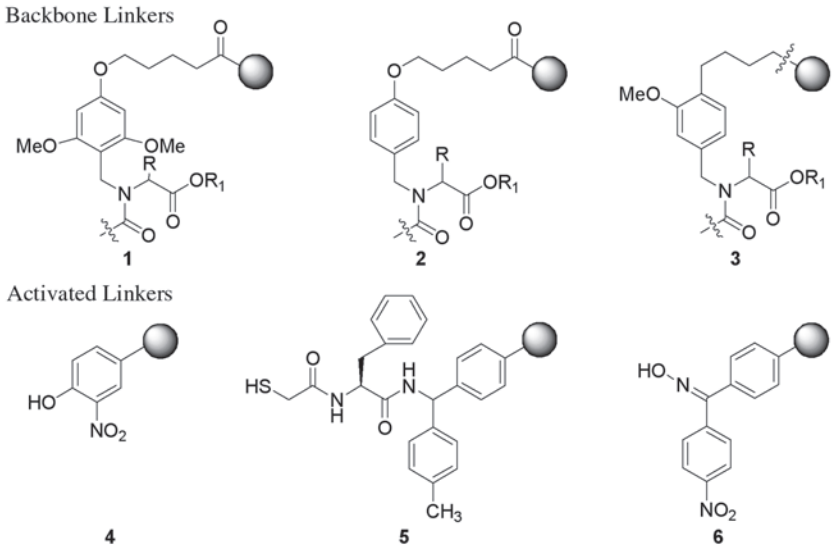


Fig. 3. Several linkers used for cyclization. For backbone linkers 1 and 2 small libraries of cyclic peptides have been synthesized *(26–28)*. Activated linkers 4 to 6 have also been used in cyclic peptide libraries *(29–32)*.

Fig. 3 *(29–32)*. A minor disadvantage of this method is that it produces protected cyclized peptide, which must be purified, deprotected, and then purified again. The backbone amide and side-chain attachment strategies alleviate this, as synthesis on resin is followed by a combined deprotection/cleavage step. However, these strategies produce the cyclic peptide in solution as a complex mixture with scavengers and cleaved protecting groups (**Fig. 4**). As a consequence, all these methods require a significant effort to be spent on purifying the cyclic peptide after synthesis.

A further extension of the activated linker principle is known as the "safety-catch" linker *(1)*. The safety-catch linker provides the capacity to remove the protecting groups on resin before cleavage by cyclization, alleviating the need for further workup. This is a very attractive process when producing large libraries. The only disadvantage of this strategy is that the side-chain functionalities are deprotected before cyclization, and thus side-chain cyclized products may be observed.

1.3. Safety-Catch Linker

There are many examples in the literature describing "safety-catch" linkers *(1)*. The basic principle, is well illustrated with linker **7** (**Fig. 4**) *(33)*. The synthesis begins with the addition of an amino acid as the symmetrical anhydride

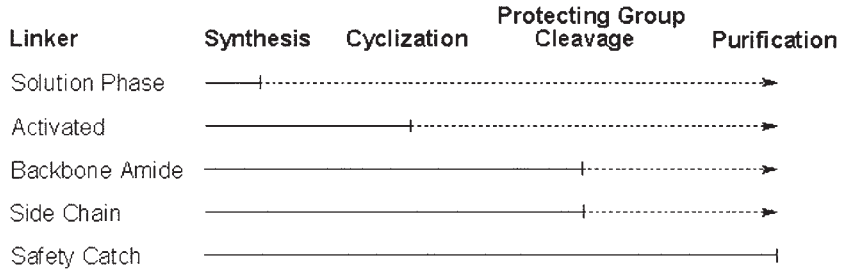


Fig. 4. The extent to which cyclic peptide synthetic strategies may be accomplished on solid support. Note that the catechol safety-catch linker requires the inversion of the order of synthesis; where the protecting groups are removed followed by cyclization.

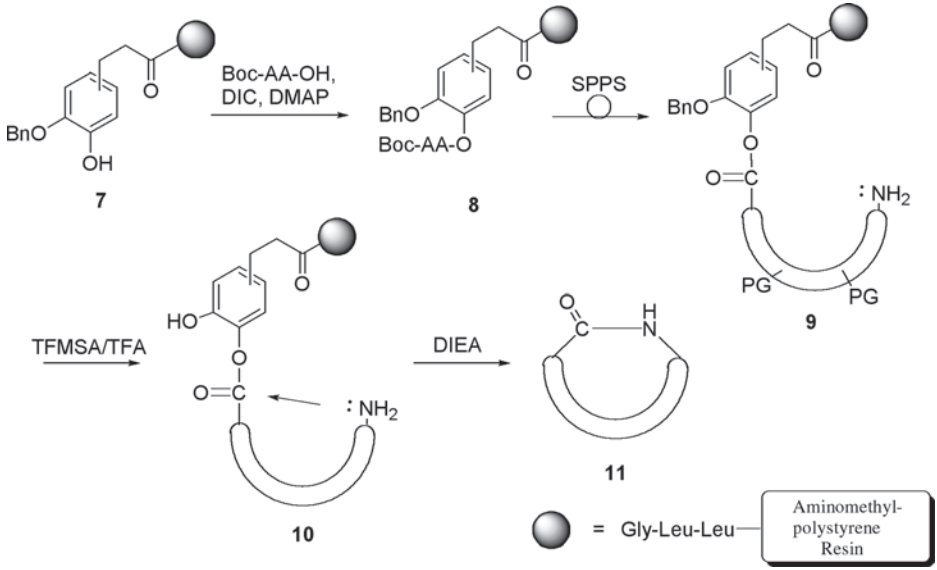


Fig. 5. A "safety-catch" linker strategy for cyclic peptide synthesis.

to the protected catechol on resin **7** using diisopropylcarbodiimide (DIC) to form the symmetrical anhydride and dimethylaminopyridine (DMAP) to deprotonate the catechol (*see* **Note 2**). This was followed by standard BOC-solid-phase peptide synthesis until the entire peptide chain **9** was assembled (*see* **Note 3**). Once complete, all protecting groups including the benzyl on the linker were removed using standard peptide synthesis cleavage conditions (TFMSA/TFA or TFMSA/DCM) (*see* **Notes 4** and **5**). Loss of the benzyl group on the catechol provided a significant rate enhancement to nucleophilic attack on the C-terminal of the peptide, so that upon treatment with diisopropylethylamine (DIEA), the peptide cyclized forming the desired product **11** (*see* **Note 6**). **Figure 5** describes

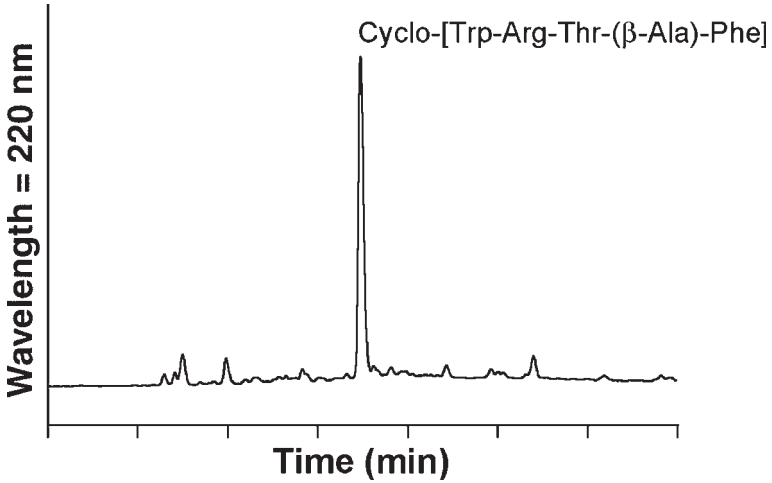


Fig. 6. Crude HPLC Cyclo-[Trp-Arg-Phe-Gly-(β-Ala)-Phe] synthesized using BOC-chemistry and the safety-catch linker. During synthesis BOC-Trp and BOC-Thr side chains were left unprotected while Pbf was used for the orthogonal protection for the Arg residue.

a crude HPLC profile of the cyclic pentapeptide Cyclo-[Trp-Arg-Phe-Gly-(β-Ala)-Phe] synthesized using the safety-catch linker.

1.4. Automation for Library Synthesis

To produce large libraries of cyclic peptides two possible strategies can be applied: discrete and split-and-pool synthesis. For our purposes (>2000 cyclic peptides), we decided to apply the split-and-pool approach coupled with a directed sorting method ***(34)***. Such an approach combines the efficiencies of split-and-pool synthesis with the highly desirable isolation of discrete compounds ***(35, 36)***. When synthesizing large libraries of cyclic peptides using a directed sorting approach, several factors can significantly affect efficiencies and cost of library production. The most obvious of these is the cyclic redundancy inherent in cyclic peptides. For example, the cyclic peptide YRFGD is the same as the cyclic peptide RFGDY, but, due to the directionality of the amide bond, this is not the same as cyclic YDGFR. It is important to remove such redundancies when synthesizing libraries of cyclic peptides. For example, the maximum number of possible combinations of cyclic pentapeptides comprising all L natural amino acids is 3.2 million sequences; eliminating cyclic redundancies reduces this to 632,016 unique cyclic peptides.

In addition, efficiencies can also be attained when considering the synthetic process used to make the large library of cyclic peptides. In general, large librar-

ies will be assembled in a series of "batches," where a batch is a number of compounds (a sublibrary) that can be efficiently handled in the laboratory. For example, 2000 compounds can be made in nine batches of approx 240 compounds. Significant gains in synthetic efficiency can be attained by deciding what compounds go into what batch (sublibrary), particularly when undertaking a directed sorted synthesis as outlined here. If different peptides have the same amino acid in the same position, then coupling that amino acid is achieved using a common reaction vessel. By rotating the peptide sequence it is possible to maximize the number of aligned amino acids. This means fewer coupling reactions are required, reducing the time taken to synthesize large libraries of cyclic peptides. At each position of the cyclic peptide, the total number of reaction vessels for amino acid coupling is therefore the total number of unique amino acids in that position.

The efficiencies gained by rotating the peptides is termed *sequencing*, and deciding into which sublibrary they should be placed is termed *batching*. The *schedule* is the list of sequences to be synthesized in each batch. Scheduling algorithms that allow consideration of in which sequence order a given peptide should be synthesized, and to what batch (sublibrary) the peptide should be added, are difficult to develop. This is a consequence of allowing both batching and sequencing to change simultaneously, which results in a vast number of possible schedules—all permutations of batches and of peptide rotations (sequencing) are possible, so the search space of all schedules is combinatorially very large.

The gradient descent algorithm ***(37)*** optimizes both sequencing and batching and significantly minimizes the number of coupling reactions. We usually target approx 2000 unique peptides with a batch size of no more than 240. We decided on 240 sequences per batch, as it was chemically practical. This gradient descent algorithm enables the chemist to greatly reduce the total number of amino acid couplings: 1797 pentapeptide sequences go from 1968 couplings to 407 (saving 79.3%) and 1954 hexapeptide sequences go from 2458 coupling reactions to 587 (saving 76.1%).

Although this optimization of schedules decreases coupling steps significantly, there are several chemical limitations of the safety-catch linker, which must be included into the optimization algorithm. The main limitations arise due to the requirement of positioning basic residues (for example, Arg and Lys) away from the C-terminus. These amino acids have the capacity of cyclizing through the side chain. This was primarily observed with Arg-containing peptides where δ-lactam formation can predominate. The side-reaction is commonly observed upon activation of unprotected Arg in normal peptide synthesis; since the activated safety-catch linker is effectively an activated ester, it is not surprising that this side product was observed. The side-reaction can be easily avoided by simply moving the Arg to a more distant position from the linker.

Presumably, this is a consequence of reducing the effective concentration of the guanido group, by distancing the Arg side chain from the active ester, and may also be because of the size of the ring being formed. Positioning the Arg away from the C-terminus allows the more nucleophilic N-terminus to compete. Similar issues occurred with other basic residues. For example, lysine can undergo side-chain to head cyclization, this can be overcome by FMOC-protection of the side chain amine, which is subsequently removed after cyclization.

1.5. Protecting Groups for BOC-Amino Acids

The central problem that may arise when using the catechol safety-catch linker is during the activation/deprotection step (*see* **Notes 4** and **5**). The best results were obtained when using HF but for large libraries this procedure is not practical. As such the optimal cleavage mixture was found to be a TFMSA/TFA or TFMSA/DCM solution with *p*-cresol or triisopropylsilane (TIPS) as scavenger. When using this mixture a recurring issue is orthogonal protection of the amino acids. The selection of protecting groups is paramount for successful synthesis. **Table 1** lists the amino acids and the corresponding protecting groups for the 20 AAs we have commonly employed for synthesis.

2. Materials

2.1. Equipment

In the laboratory we have several automatic and semiautomatic robotic apparatus including the IRORI Nanokan optical encoding system for split- and pool-synthesis ***(34)***; and synthesizers from Advanced ChemTech and the MiniBlock™ from Mettler, Toledo, OH.

2.2. Reagents

1. *Tert*-butoxycarbonyl (BOC)-amino acids.
2. Peptide synthesis grade dimethylformamide (DMF).
3. Trifluoroacetic acid (TFA).
4. Diispropylcarbodiimide (DIC).
5. Diisopropylethylamine (DIEA).
6. Dimethylaminopyridine (DMAP).
7. 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyl-uronium hexafluorophosphate (HBTU).
8. Anhydrous hydrogen fluoride (HF).
9. Triisopropylsilane (TIPS).
10. Solvents: AR-grade ethyl acetate, methanol, dichloromethane, chloroform, hexane, and acetone; HPLC-grade acetonitrile.
11. Aminomethylpolystyrene resins with a substitution value in the range of 0.26 to 0.41 mmol/g.

Table 1
Summary of Side-Chain Protecting Groups on Amino Acids Compatible With Catechol Safety-Catch Linker During Large Library Synthesis (More Than 2000 Compounds)

BOC-AA-OH	Issues and comments
Arg(Pbf)	Although Mts has been used successfully for Arg protection, for large libraries we sometimes observed incomplete deprotection. To circumvent this problem the Mts protecting group was replaced with Pbf. Removal occurs at the same time as BOC- deprotection during peptide synthesis.
Asp(Bzl) Glu(Bzl) Tyr(Bzl)	For acidic residues and Tyr, the Bzl protecting group is preferred. The cyclohexyl protecting group for Asp and Glu (preferred for HF cleavage conditions) was not removed under 10% TFMSA/TFA conditions, although at 40% TFMSA/TFA the cyclohexyl group was removed.
His(Boc)	The histidine was BOC-protected so that removal can be performed during peptide synthesis. The Tosyl protection can also be use for problematic sequences during synthesis.
Ser, Thr, Trp, Gln	Ser, Thr, and Trp were left unprotected for short sequences (i.e., 5 to 6 amino acids in length). Initially we used benzyl-protecting groups for the alcohols, but resulting from incomplete deprotection for some libraries, we left the side-chain unprotected. For Trp, alkylations during activation/deprotection step was sometimes observed. As such we sometimes include the formyl protecting group on the indole for difficult sequences.
Asn(Trityl), Cys(4-MeBzl) Lys(ClZ),	Standard amino acid protecting groups that were used routinely in BOC-peptide synthesis. For Lys the ε-nitrogen has the potential to cyclize. We sometimes protect this with the fluorenylmethyloxycarbonyl (FMOC) protecting group. A further FMOC deprotection step (piperidine:DMF, 1:1) is then required after cyclization to obtain unprotected cyclic peptide.
Ala, Gly, Ile, Leu, Met, Phe, Pro, Val	Standard amino acids that are used in peptide chemistry.

3. Methods

3.1. Attachment of Safety-Catch Linker to Resin

3-(1-hydroxy-2-benzyloxyphenyl)propanoic acid 13.6 g (50 mmol) and 18.9 g (50 mmol) HBTU was dissolved in 300 mL DMF. To this was added 10.75 mL (61 mmol) DIEA and the resulting solution added to 61 g (0.41 mmol/g) pre-coupled H-Gly-Leu-Leu-aminomethylpolystyrene. The resin was shaken over-

night and until >99% completion as determined by a quantitative ninhydrin test. The eluent was filtered off and the resin washed with DMF (4 × 50 mL).

3.2. Acylation of Safety-Catch Linker

The symmetric anhydrides of BOC-amino acids were prepared by adding 1.565 mL (10 mmol) DIC to a stirred solution of 15 mmol BOC-amino acids in 10 mL DCM. The solution was then added to H-Gly-Leu-Leu-aminomethylpolystyrene resin previously derivatized with the linker **7**. Solid DMAP (approx 50 mg) was then added to the solution and shaking continued for 3 h. The resin was then filtered and washed with DMF (4 × 12 mL) and then DCM (4 × 12 mL).

3.3. Resin Preparation for Automated Synthesis of Linear Peptides

Approximately 100 mg of resin was distributed to each of the reaction block wells (of an ACT block or a Bohdan block) by pipetting a slurry of the resin in DMF/DCM (3:1) or as dry resin into each IRORI kan. The peptides were then assembled by the combinatorial chemistry apparatus suited for parallel or split- and pool-synthesis ***(34)*** using *in situ* neutralization/HBTU activation protocols for BOC chemistry. The resin was initially washed with DCM and the BOC protecting group removed by washing twice with a 40% solution of TFA in DCM.

3.4. Automated Synthesis of Linear Peptides for Discrete Libraries

To each well containing the preloaded resin was added a 0.3 mmol solution of the appropriate amino acid and 142.8 mg (0.3 mmol) HBTU in 1 mL DMF. DIEA 50 µL (0.3 mmol) was then added to each reaction mixture. The reaction block was shaken for 1 h, emptied, and the resin washed with DMF (3 × 1 mL). Deprotection, coupling and washing cycles were repeated stepwise to assemble the linear peptides.

3.5. Automated Synthesis of Linear Peptides Using Split- and Pool-Synthesis (34)

To each vessel containing a set number of kans was added a solution of 0.25 mmol of the appropriate amino acid, 0.25 mmol HBTU, and 0.25 mmol DIEA in 1 mL DMF. The vessel was shaken for 1 h and emptied before a second coupling was initiated. The resin was flow-washed with DMF and DCM before the kans were regrouped and treated with 40% TFA/DCM solution. Coupling, deprotection, and washing cycles were repeated stepwise to assemble the linear peptides.

3.6. Activation/Deprotection and Cyclisation of Linear Peptides

After the final BOC-deprotection of the linear peptide the resin was flow-washed with a 2% solution of DIEA/DMF followed by DCM and dried under

nitrogen for 3 h. For activation of discrete libraries, 100 µL TIPS was added per 100 mg of resin in each reaction block well. A solution of TFMSA:DCM (2:8) was prepared by the dropwise addition of TFMSA to a vigorously stirred solution of DCM at −5°C, which was subsequently allowed to warm to room temperature. 1 mL TFMSA/DCM was then added to each of the reaction block wells and the resin shaken for 2 h. For split- and pool-synthesis ***(34)***, the resin-filled kans were coated with scavenger (TIPS) via vigorous mixing in a large round-bottom flask with a large magnetic stirrer for 30 min. DCM was then added to the flask and mixing was continued for an additional 30 min. TFMSA was slowly added over 10 min and the activation/deprotection was carried out for 1 to 2 h before filtering. After activation/deprotection the resin was washed with DCM and DMF. A solution of DIEA (5%) in DMF was added to each of the reaction wells and the reaction block was shaken overnight or up to 7 d (*see* **Note 6**). We commonly employ a second cyclization with a second 1-mL aliquot of 5% DIEA/DMF solution. The eluants were combined from the reaction block into a deep-well titer plate, the resin was rinsed with DMF, and the eluent was collected into a separate titer plate. The solvent was removed by centrifuge evaporation and the product dissolved in 0.1% TFA in 45% CH_3CN:55% H_2O. Further drying resulted in crude cyclic peptides. Each cyclic peptide was then purified using mass directed fractionation yielding 1 to 5 mg of cyclic peptides depending on cyclisation efficiencies.

4. Notes

1. There are five commonly observed classes of cyclic peptides. The most common of these is a head-to-tail cyclic product, seen in **Fig. 2**, in which lactamization occurs between the carboxyl and amino termini. Alternatively, cyclization can be effected between a side chain and the carboxyl terminus or amino terminus of the peptide. Cyclization may also be achieved between two side chains, which often involves the use of an additional chemical spacer. The final method by which cyclization may be attained is through two backbone amide nitrogens. All these strategies have different requirements for the orthogonal protecting groups employed.
2. When synthesizing cyclic peptides a low substitution resin was required to impede possible oligomerization. We tend to use resin with low substitution values between 0.21 to 0.45 mmol/g. We also include a spacer group usually [Leu-Leu-Gly] between the resin and the safety-catch linker. Previously, we have noted that peptide assembly is more effective when this spacer group is present ***(29)***.
3. Synthesis of the linear peptides on resin was accomplished with 1-h couplings with strong agitation.
4. The most problematic issue using the catechol safety-catch linker is the activation/deprotection steps. Although best results are obtained with HF, for large libraries in our laboratory this was not possible. Therefore we tend to use TFMSA/DCM as the activation mixture.

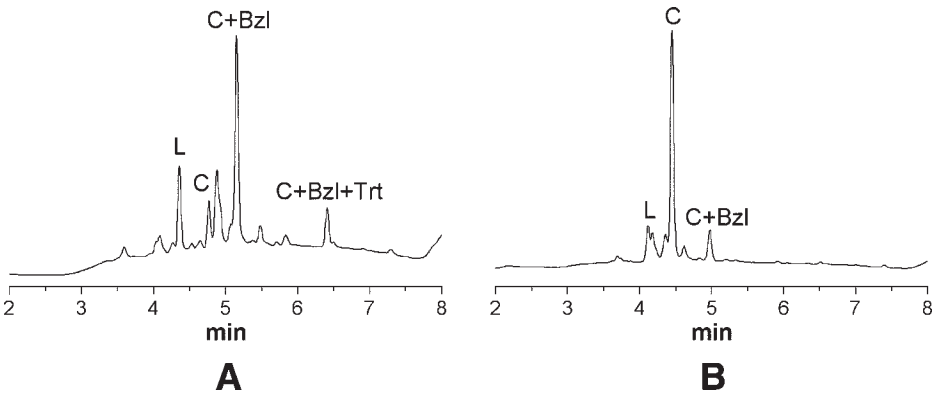


Fig. 7. Crude HPLC of Cyclo-[Phe-Asn-Val-(D)Glu-Ala-Gly]. HPLC A after one treatment of TFMSA/TFA/p-cresol. HPLC B after two treatments of TFMSA/TFA/p-cresol. C = Cyclic peptide, Bzl = benzyl, Trt = Trityl, L = Linear.

a. TFA adducts are often present when synthesizing cyclic peptides containing serine and threonine. This side-reaction can be attributed to acylation of the alcohol side-chain followed by an $O \rightarrow N$ acyl shift to the terminal amine. This side-reaction was first described by Hübener et al., who observed it in N-terminal threonine sequences under conditions of 99% aqueous TFA ***(38)***. They proposed a mechanism that proceeded through an *O*-trifluoroacetylated intermediate followed by intrachain $O \rightarrow N$ acyl shift or interchain aminolysis. In our case the N-trifluoroacetyl side-product effectively caps the N-terminus, eliminating the potential to cyclize. The initial serine *O*-trifluoroacetylation is thought to occur during TFMSA/TFA side-chain deprotection. Hence, this side-product was greatly reduced by replacing the TFA with DCM as the solvent in the deprotection mixture and further by the neutralization of the resin after final BOC-deprotection to remove any TFA salts.
b. High acid treatment often results in alkylations (especially in the presence of tryptophan residues) and/or incomplete activation deprotection. An example is shown in **Fig. 7** where incomplete activation/deprotection has occurred after one treatment. Because of the potential for inconsistent deprotection, we regularly deprotect and activate the linker with two treatments for two hours at 10% TFMSA/DCM, or one treatment at 20% TFMSA/DCM.

5. Aspartimide formation is sometimes a significant side-reaction in sequences containing Asp-Gly, Asp-Ser, and Asp-Ala residues. This side-product can be easily hydrolyzed to give a mixture of α- and β-peptides or, if this is deemed undesirable, then moving the aspartic acid to the C-terminal position can easily prevent it, as aspartimide formation is not possible with Asp or Glu at the C-terminal posi-

tion. Importantly, side-chain cyclization of Asp or Glu to the N-terminus is not possible, as the only activated ester is the C-terminal carboxyl group that is attached to the linker.

6. Cyclization is performed using a mixture of dry 2% to 5% DIEA/DMF solution. It is vitally important for the procedure that this mixture is free of water. Therefore, we use a large glove bag or an inert atmosphere chamber. This enables cyclisation to be easily performed from 1 to 7 d without hydrolysis being a byproduct.

Acknowledgments

The authors would also like to thank present and past group members for their contributions to the concepts presented, particularly Simon Golding and Wim Meutermans. Various colleagues also contributed, particularly Peter Cassidy, Alun Jones, Paul Alewood, Garland Marshall, and Peter Andrews. This work was supported by Novo Nordisk GlaxoSmithKline, the Australian Research Council and the ARC Special Research Centre for Functional and Applied Genomics.

References

1. Horton, D. A., Bourne, G. T., and Smythe, M. L. (2002) Exploring privileged structures: the combinatorial synthesis of cyclic peptides. *J. Comp-Aided Mol. Des.* **16,** 415–430.
2. MacDonald, M. and Aube, J. (2001) Approaches to cyclic peptide β-turn mimics. *Curr. Org. Chem.* **5,** 417–438.
3. Emmel, E. A., Verweij, C. L., Durand, D. B., Higgins, K. M., Lacy, E., and Crabtree, G. R. (1989) Cyclosporine-A specifically inhibits function of nuclear proteins involved in T-cell activation. *Science* **246,** 1617–1620.
4. Mootz, H. D. and Marahiel, M. A. (1997) The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179,** 6843–6850.
5. Kratzschmar, J., Krause, M., and Marahiel, M. A. (1989) Gramicidin-S biosynthesis operon containing the structural genes GRSA and GRSB has an open reading frame encoding a protein homologous to fatty-acid thioesterases. *J. Bacteriol.* **171,** 5422–5429.
6. Sanchez-Quesada, J., Ghadiri, M. R., Bayley, H., and Braha, O. (2000) Cyclic peptides as molecular adapters for a pore-forming protein. *J. Am. Chem. Soc.* **122,** 11,757–11,766.
7. Hirschmann, R., Yao, W. Q., Cascieri, M. A., et al. (1996) Synthesis of potent cyclic hexapeptide NK-1 antagonists. Use of a minilibrary in transforming a peptidal somatostatin receptor ligand into an NK-1 receptor ligand via a polyvalent peptidomimetic. *J. Med. Chem.* **39,** 2441–2448.
8. Rohrer, S. P. and Schaeffer, J. M. (2000) Identification and characterization of subtype selective somatostatin receptor agonists. *J. Physiol. Paris* **94,** 211–215.

9. Quartara, L., Pavone, V., Pedone, C., Lombardi, A., Renzetti, A. R., and Maggi, C. A. (1996) A review of the design, synthesis and biological activity of the bicyclic hexapeptide tachykinin NK2 antagonist MEN 10627. *Regul. Pept.* **65,** 55–59.
10. Schiller, P. W., Berezowska, I., Nguyen, T. M.-D., et al. (2000) Novel ligands lacking a positive charge for the δ- and μ-opioid receptors. *J. Med. Chem.* **43,** 551–559.
11. Boer, J., Gottschling, D., Schuster, A., Semmrich, M., Holzmann, B., and Kessler, H. (2001) Design and synthesis of potent and selective $\alpha_4\beta_7$ integrin antagonists. *J. Med. Chem.* **44,** 2586–2592.
12. Yoshida, M., Furumai, R., Nishiyama, M., Komatsu, Y., Nishino, N., and Horinouchi, S. (2001) Histone deacetylase as a new target for cancer chemotherapy. *Cancer Chemother. Pharmacol.* **48(Suppl 1),** S20–S26.
13. Yoshida, M., Horinouchi, S., and Beppu, T. (1995) Trichostatin A and trapoxin: Novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17,** 423–430.
14. Tkacz, J. S. and DiDomenico, B. (2001) Antifungals: what's in the pipeline. *Curr. Opin. Microbiol.* **4,** 540–545.
15. Okada, Y. (2001) Synthesis of peptides by solution methods. *Curr. Org. Chem.* **5,** 1–43.
16. Lee, J. H., Griffin, J. H., and Nicas, T. I. (1996) Solid-phase total synthesis of bacitracin A. *J. Org. Chem.* **61,** 3983–3986.
17. Kates, S. A., Sole, N. A., Johnson, C. R., Hudson, D., Barany, G., and Albericio, F. (1993) A novel, convenient, three-dimensional orthogonal strategy for solid-phase synthesis of cyclic peptides. *Tetrahedron Lett.* **34,** 1549–1552.
18. Tromelin, A., Fulachier, M.-H., Mourier, G., and Ménez, A. (1992) Solid-phase synthesis of a cyclic peptide derived from a curaremimetic toxin. *Tetrahedron Lett.* **33,** 5197–5200.
19. Cabrele, C., Langer, M., and Beck-Sickinger, A. G. (1999) Amino acid side chain attachment approach and its application to the synthesis of tyrosine-containing cyclic peptides. *J. Org. Chem.* **64,** 4353–4361.
20. Spatola, A. F., Crozet, Y., deWit, D., and Yanagisawa, M. (1996) Rediscovering an endothelin antagonist (BQ-123): A self-deconvoluting cyclic pentapeptide library. *J. Med. Chem.* **39,** 3842–3846.
21. Rovero, P., Quartara, L., and Fabbri, G. (1991) Synthesis of cyclic-peptides on solid support. *Tetrahedron Lett.* **32,** 2639–2642.
22. Spatola, A. F., Darlak, K., and Romanovskis, P. (1996) An approach to cyclic peptide libraries: reducing epimerization in medium sized rings during solid-phase synthesis. *Tetrahedron Lett.* **37,** 591–594.
23. Trzeciak, A. and Bannwarth, W. (1992) Synthesis of 'head-to-tail' cyclised peptides on solid support by FMOC chemistry. *Tetrahedron Lett.* **33,** 4557–4560.
24. McMurray, J. S. (1991) Solid-phase synthesis of a cyclic peptide using FMOC chemistry. *Tetrahedron Lett.* **32,** 7679–7682.
25. McMurray, J. S., Lewis, C. A., and Obeyesekere, N. U. (1994) Influence of solid support, solvent and coupling reagent in the head-to-tail cyclisation of resin-bound peptides. *Pept. Res.* **7,** 195–206.

26. Jensen, K. L., Alsina, J., Songster, M. F., Vágner, J., Albericio, F., and Barany, G. (1998) Backbone amide linker (BAL) strategy for solid-phase synthesis of C-terminal-modified and cyclic peptides. *J. Am. Chem. Soc.* **120,** 5441–5452.
27. Bourne, G. T., Meutermans, W. D. F., Alewood, P. F., et al. (1999) A backbone linker for BOC-based peptide synthesis and on-resin cyclisation: synthesis of stylostatin 1. *J. Org. Chem.* **64,** 3095–3101.
28. Gu, W. and Silverman, R. B. (2003) New stable backbone linker resins for solid-phase peptide synthesis. *Org. Lett.* **5,** 415–418.
29. Fridkin, M., Patchornik, A., and Katchalski, E. (1968) Use of polymers as chemical reagents II. Synthesis of bradykinin. *J. Am. Chem. Soc.* **90,** 2953–2957.
30. Richter, L. S., Tom, J. Y. K., and Burnier, J. P. (1994) Peptide-cyclizations on solid support—a fast and efficient route to small cyclopeptides. *Tetrahedron Lett.* **35,** 5547–5550.
31. DeGrado, W. F. and Kaiser, E. T. (1982) Solid phase synthesis of protected peptides on a polymer bound oxime: preparation of segments comprising the sequence of a cytotoxic 26-peptide analog. *J. Org. Chem.* **47,** 3258–3261.
32. DeGrado, W. F. and Kaiser, E. T. (1980) Polymer-bound oxime esters as supports for solid-phase peptide synthesis. The preparation of protected peptide fragments. *J. Org. Chem.* **45,** 1295–1300.
33. Bourne, G. T., Golding, S. W., McGeary, R. P., et al. (2001) The development and application of a novel safety-catch linker for BOC-based assembly of libraries of cyclic peptides. *J. Org. Chem.* **66,** 7706–7713.
34. Nicolaou, K. C., Pfefferkorn, J. A., Mitchell, H. J., et al. (2000) Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10,000-membered benzopyran library by directed split-and-pool chemistry using NanoKans and optical encoding. *J. Am. Chem. Soc.* **122,** 9954–9967.
35. Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A., and Gallop, M. A. (1994) Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. *J. Med. Chem.* **37,** 1385–1401.
36. Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A., and Gordon, E. M. (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* **37,** 1233–1251.
37. Darwen, P. J., Bourne, G. T., Nielson, J., Tran, T. T., and Smythe, M. L. (2003) A gradient descent algorithm for minimizing the number of steps required for synthesis of cyclic-peptide libraries. *Personal communication.*
38. Hübener, G., Goehring, W., Musiol, H. J., and Moroder, L. (1992) N-α-trifluoroacetylation of N-terminal hydroxy amino acids: a new side reaction in peptide synthesis. *Peptide Res.* **5,** 287–292.

11

High-Throughput Peptide Synthesis

Michal Lebl and John Hachmann

Summary

The methodologies of high-throughput peptide synthesis are overviewed and discussed. Particular focus is given to the techniques applicable to laboratories with a limited budget. Automated solutions for synthetic problems are also discussed.

Key Words: Automation; parallel synthesis; solid phase synthesis; manual synthesizer; centrifugation; review.

1. Introduction

According to the Protein Biochips 2003 Report, the protein biochip market is underserved, despite growing from $70 million in 2001 to $100 million in 2002. The report projects that the market will grow to more than $400 million in 2007, a compound annual growth rate of more than 35%. This need for peptide synthesis technology is being driven by the increasing availability of sequence information. The initial draft of the human genome sequence has been finished and results were published recently ***(1,2)***. With this explosion in sequence information has come the development of new technologies that take advantage of this knowledge. Some of these technologies have a high dependence on the availability of peptides—for example, array-based technologies that allow the parallel analysis and quantitation of thousands of proteins at a time.

The current worldwide supply for custom-made peptides is estimated to be around $10 billion (this amount includes pharmaceutically important peptides). Principal components of the market are peptides for studies of protein–protein interactions, finding antibody epitopes, and analogs of biologically active peptides and potential drugs. Added to this market is a significant business in peptide synthesizers, reagents, and supports. Large-scale proteomics research is

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

Table 1
Leading Suppliers of Custom Peptides

Company	Country	Price ($/amino acid residue)
Peptron	Daejeon, South Korea	10
New England Peptide	Gardner, MA, USA	15
Genemed Synthesis	South San Francisco, CA, USA	15
American Peptide Company	Sunnyvale, CA, USA	16
Invitrogen	Carlsbad, CA, USA	18
Sigma/Genosys	The Woodslands, TX, USA	18
AnyGen	Kwang-ju, Korea	20
AnaSpec	San Jose, CA, USA	25

not yet a primary driver of this growth. One particular area that will stimulate the growth of the peptide market is the isotope-coded affinity tagging (ICAT) approach to protein profiling ***(3,4)***. While there are many different methods for performing protein profiling, none of them is as efficient and multiplexable as ICAT. To unequivocally identify and quantify a particular protein, two or three specific peptides are required. With an estimated 500,000 common proteins present in the human proteome, just one application will require 1.5 million peptides. The present cost of custom peptide synthesis (*see* **Table 1**) precludes the wide application of this technology. At $15 per amino acid residue, the cost of necessary peptides would be approx $225 million. As technologies capable of bringing the cost of custom peptides to a level of $0.5 per amino acid residue, or lower (i.e., 30 times lower than current running costs) are becoming available, ICAT studies of the whole proteome become feasible.

As evident from previous paragraphs, peptide synthesis technology is of major strategic importance in the field of proteomics. Currently, there are various instruments for automatic or even parallel synthesis of peptides. While these technologies meet the modest requirements of most experiments today, they are inadequate for the manufacturing needs looming in the very near future. Current synthesis technologies do not meet the need for the cost-effective manufacture of large numbers of peptides (tens of thousands to millions of sequences).

1.1. Parallel Synthesis

Merrifield's idea of the synthesis of peptides on solid support ***(5–9)*** fundamentally changed the thinking of peptide chemists (even though, depending on the personality of the scientists, the adoption of the solid phase methodology sometimes required a very long "induction period"). Later, as it became

apparent that the transformation of sometimes unpredictable behavior of synthetic intermediates into predictable behavior of solid support with attached organic moieties makes synthetic process and purification of intermediates very simple, organic chemists worked hard to show that almost all types of organic transformations can be performed on solid support and solid phase synthesis was broadly embraced by chemists in all fields.

Solid supports allow an easy realization of parallel synthesis. It requires just the compartmentalization of solid phase and the individualized delivery of reactants. This concept of compartmentalization was later pushed to its limit by the realization that each particle of solid support can serve as an individual compartment in which individual peptides can be produced to create libraries of millions of peptides or any other molecular species ***(10)***. This one-bead–one-compound concept is covered in detail elsewhere ***(10,11)*** and will not be discussed here in detail.

2. Methods

2.1. Manual Approaches to High-Throughput Synthesis

The requirements of epitope mapping were addressed by the synthesis of peptide arrays on polyethylene "pins." This technique, pioneered by Geysen and colleagues ***(12–14),*** utilizes solid support in the form of rods functionalized with a layer of "synthesis friendly"–swellable polymer, on which the repetitive coupling reaction is performed. These pins are arranged in a grid mapping the microtiter plate format and coupling is realized by dipping this grid into the plate preloaded with appropriate activated amino acids. Washing and the deprotection reaction do not require segregation of individual pins, and can be realized by simple dipping of the array into the common container. The problem of limited amounts of peptide prepared on individual pins was overcome by the introduction of "crowns," or "lanterns" having higher surface areas available for synthesis ***(15)***. The original paper described the utilization of the pin-bound peptide for the evaluation of the antibody binding, but the release of peptides into solution became more popular, enabling peptides to be used for other analytical evaluations.

The manual aspect of pin synthesis is simplified by the utilization of a computer-driven light box ***(16)*** indicating that amino acid should be pipetted into which position by lighting up LEDs under appropriate wells. A very simple and helpful technique is overlaying the microtiter plate with a grid preprinted with the amino acid labels. Delivery of each building block is verified by puncturing the paper overlay, thus preventing skipping the well or delivering multiple doses into each well. Further simplification of the delivery of appropriate building blocks can be achieved by application of pipetted robots. A simple

Excel macro generating pipetting tables from a list of sequences is available online (www.5z.com/mlebl/macros.html). The area of grafted solid supports in peptide synthesis was recently reviewed ***(17)***.

Another extremely flexible parallel synthesis approach was developed by Houghten et al. ***(18,19)***. This so-called "tea-bag" synthesis utilizes the compartmentalization of polystyrene-based solid support into individual polypropylene mesh bags, which are labeled by either alphanumeric or bar codes ***(20)***, or into which a radiofrequency tag is added ***(21,22)***. The bags are re-sorted prior to each coupling reaction and all bags requiring the same amino acid are placed into the same reaction vessel. After completion of amino acid coupling, all tea-bags are combined into a larger vessel and washing and common reactions (e.g., removal of amino terminal protecting group) are performed simultaneously. The size of individual bags ranges from 25 mg to several grams. Manual sorting allows batches of up to several hundred peptides to be synthesized in parallel. Automated sorting of "NanoKans" (permeable microreactors in the form of tiny cans, filled with the resin and labeled by two-dimensional bar codes; (*see* http://www.irori.com/Products/irori_tech_nanokan.html) enabled the synthesis of thousands of compounds in parallel. This technology is marketed by Discovery Partners International (http://www.discoverypartners.com).

Parallel synthesis in tea-bags would not be as valuable if there was no method for the parallel cleavage of products from the solid support. For this reason, a 24-vessel parallel hydrofluoric acid (HF) cleavage apparatus was developed ***(23)***. Later it was found that the application of gaseous HF works well for the parallel cleavage of peptides synthesized in both tea-bags and microtiter plates ***(24,25)***. Chambers for the cleavage of 1728 products from resin in 18 microtiter plates were constructed ***(26)***.

A similar approach utilizing cellulose paper ***(27,28)*** (originally developed for DNA synthesis ***[29]***), was later modified by using a printer to label individual pieces of synthetic support ***(30)***. Cotton, as the most pure form of cellulose, was found to be a convenient support for parallel synthesis as well ***(31–34)***, and was used for the synthesis of combinatorial libraries with guaranteed uniform representation of each structure ***(35)***.

Cellulose paper is ideal for the so-called SPOT synthesis ***(36–38)*** (for an extensive review of SPOT techniques *see* **refs. *39,40***). This technique is based on the concept of inclusion volume coupling ***(41)***. Reagents are spotted onto the porous material (cellulose paper, cotton) and coupling occurs utilizing only the volume of reagent soaked into the pores of the support. No external volume is needed for the completion of the reaction. Common steps (washing, deprotection) are performed by immersion or flooding the support with the appropriate solvent. The process was later automated; a synthesizer using this technology can be purchased from Intavis AG, formerly Abimed (http://www.intavis.com/

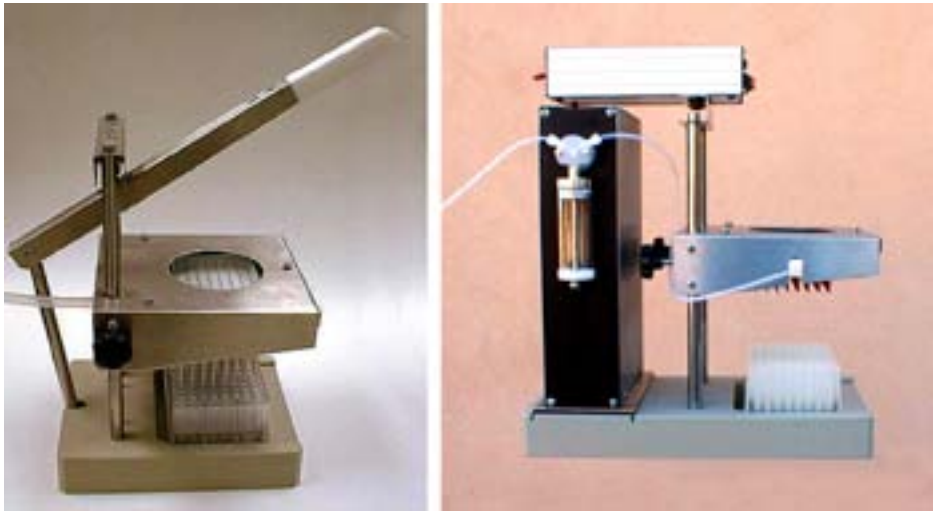


Fig. 1. Aspiration (left) and delivery (right) station for 96-well plate-based synthesis employing surface suction concept. Pushing down spring-supported handle lowers the array of flat-end needles against the surface of liquid in the wells of microtiter plate. (Pictures courtesy of Viktor Krchnak.)

autospot_s.html). The success of Jerini (http://www.jerini.de) as a provider of high-throughput peptide synthesis products is also based on SPOT synthesis.

Another very efficient parallel synthetic technique utilizes the principle of surface suction. In this case liquid is removed from the reaction mixture after the solid particle is sedimented by aspiration. However, immersion of a needle directly into the liquid could disturb the system and remove sedimented particles together with the bulk of the liquid. To prevent this problem, aspiration (suction) is applied through a needle before it touches the liquid surface and the needle is then slowly lowered against the liquid surface. This allows the liquid to be "shaved" from the surface without disturbing the bulk of the liquid. Therefore, the needle can go very close to the layer of sedimented particles without removing them from the mixture. Obviously, this method requires that the solid phase sediments in the washing step. This is relatively easily achievable, because a dense solvent can always be replaced with a less dense one by evaporation and replacement. Simple dilution with solvents such as *tert*-butyl methyl ether will also achieve sedimentation of solid particles ***(42)***. Aspiration and delivery stations for 96-well microtiter plates (**Fig. 1**) are commercially available (http://www.torviq.com), as well as software to drive the syringe pump (http://www.promptpublishing.com/software.html).

The first commercial manual multiple synthesizer, RaMPS, was marketed by E. I. duPont de Nemours. "Semiautomation" in this instrument consists of the simultaneous evacuation of reactors ***(43)***. This synthesizer, however, is no longer available. For laboratories that are eager to start the parallel synthesis of peptides with minimal investment, synthesis in plastic fritted syringes can be highly recommended ***(44–49)***. The only equipment needed is plastic disposable syringes with tightly fitting frit material at the bottom, commercially available from several sources; (*see*, for example, http://www.5z.com/csps or http://www.torviq.com) and a suitable shaker (rotary, linear, etc.). The repeated use of the syringes is not recommended owing to the possibility of contamination. For the synthesis of longer peptides (more than 30 mer) it may even be advisable to replace the syringe with a new one during the synthesis to prevent mechanical failure. The syringe is charged with the appropriate amount of the resin, taking into an account the swelling in the solvents used in the course of the synthesis as well as the increase of the resin volume in the synthesis. All steps are simply performed by aspirating the appropriate reagent through the needle, stoppering the needle by sticking it into a rubber stopper, and placing the syringe on the shaker. Every syringe is appropriately labeled and no particular arrangement is therefore necessary. If the syringe is equipped with the bar code, than the management software simplifies the work of the chemist tremendously (http://www.promptpublishing.com/software.html). Performing the synthesis in the syringe has one additional advantage—the possibility of monitoring resin swelling, which is very indicative to the internal aggregation of the growing peptide chains and accompanying problems in coupling and deprotection ***(44–53)***. Syringes allowing utilization of extreme conditions are constructed from Teflon® (http://www.torviq.com).

Plastic syringes can be arranged in various types of "blocks." One of the classical designs is shown in **Fig. 2**. This instrument was designed by Krchnak and Vagner ***(48)*** and is constructed from Teflon®, polypropylene, glass, and stainless steel. Reactors equipped with plastic frits are either stoppered on both ends by flexibly mounted multistoppers (for coupling) or the whole block of reactors is placed on a suction table and all reactors are washed simultaneously. Individual spring-supported stoppers arranged to match the grid of syringes allows individual stoppers to be pushed onto the openings with the same force, thus compensating for any irregularities in the syringe array. After placement of this "multiblock" onto the lower multistopper, coupling reagents are added, the upper multistopper is attached, and the assembly is placed onto a shaker, or it is shaken only occasionally by hand. Each reactor can hold up to 2 mL of swollen resin, i.e., about 0.4 g of resin. Reactions can be performed in an ultrasonic bath ***(47)***. The multiblock design is optimal for the noninvasive continu-

Fig. 2. Multiblock. Arrangement of 42 syringes with multistoppers and washing station. Left—Multiblock with upper and lower multistopper attached to it for coupling reaction next to the washing station. Right—Multiblock with cover used for resin distribution. (Photo by ML.)

ous monitoring of amide bond formation ***(54)***. Unique features of this block are the capability of randomization and the synthesis of one-bead-one-compound libraries. For this purpose, the resin is combined in the common area above the syringe tops by simple inversion of the block. After shaking and inversion, sedimentation uniformly distributes the resin. Detailed instructions for use of this block are available on the internet (http://www.5z.com/csps/comer/c_synth/manual.html).

Another unique application of synthesis in plastic syringes is the concept of "domino blocks" ***(55)***. In this case, the syringes are used individually for the coupling step, but for common operations they are attached to the block, which is connected via a vacuum source to common reagents and solvents (*see* **Fig. 3**). Blocks are placed on the shaker platform and syringes are evacuated. After switching the selector valves into a proper position, liquid is aspirated into all syringes simultaneously and syringe contents are shaken before repetition of the process. This arrangement eliminates tedious "plunger pushing" and increases the productivity of syringe synthesis many times (http://www.torviq.com).

Ontogen has developed OntoBLOCK, a system that contains 96 reaction vessels and is capable of a wide array of organic syntheses. In combination with a pipetting system operating on multiple blocks, an in-house combinatorial chemistry automation system can produce 1000 to 2000 peptides or small organic molecules per day ***(56)***. Bohdan Automation Inc. (now acquired by Mettler Toledo) developed a similar reaction block (http://www.bohdan.com/miniblock.htm) that utilizes the simultaneous pinching of flexible tubing as a mechanism for the closing of vessels (*see* also **ref. *57***).

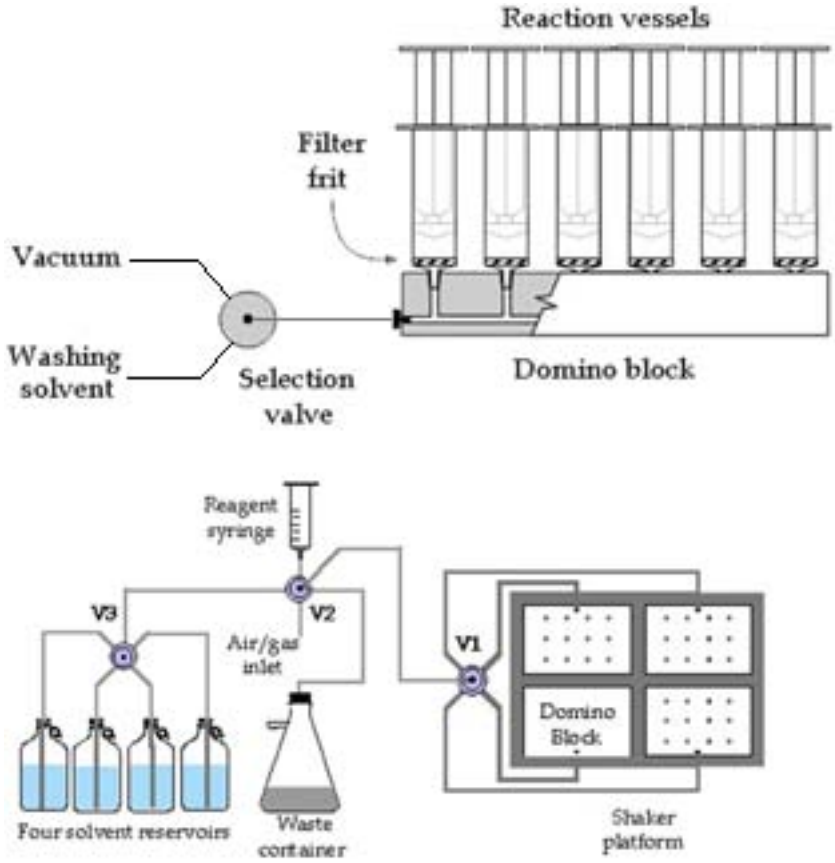


Fig. 3. Scheme of operation of "domino blocks." The block is either attached to vacuum (for emptying) or to source of liquid (for charging). Small amount of air or inert gas is introduced through V2 for creation of bubble in the syringe, allowing efficient mixing.

2.2. Synthesis Automation

Solid phase synthesis lends itself easily to automation. This fact was readily apparent to Merrifield, who built the first automated synthesizer ***(8,58,59)***. Merrifield's pioneering effort was soon followed by other laboratories ***(60–87)*** and several synthesizers appeared on the market. The development of the Fmoc synthetic strategy ***(88,89)*** allowed for the substantial simplification of automatic synthesizers, since handling of the very "unfriendly" reagent, hydrofluoric acid, was no longer necessary. Cambridge Research Biochemicals was the first to introduce a Fmoc-based synthesizer, PEPSYNthesizer, to the market. This machine was capable of only one step of the synthesis, but it applied the

revolutionary principle of monitoring the reaction by following the UV absorbance (decrease during coupling, increase during deprotection) of the reaction solution recirculated through the column packed with solid support ***(90)***. (Actually, Sheppard defined as truly automated only systems utilizing feedback control ***[65]***). The completely automatic Pharmacia-LKB synthesizer Biolynx 4170 later used a similar concept. Curiously, none of these machines is available today.

Table 2 summarizes contemporary synthesizers and companies manufacturing them and, for historical reasons, also includes some discontinued products. As you can see, synthesizers with scales covering milligram to kilogram capacity are available with automation ranging from performing only single steps without human interference up to machines incorporating feedback control based on several monitoring concepts. Very soon, however, it became apparent that serial synthesis alone would not satisfy the demand for the large-scale supplies of peptides.

In order to accelerate the synthesis of large single-compound arrays, much effort was devoted to the design of high-throughput organic synthesizers (*see* e.g., **ref. *91***). Numerous companies have recently started to develop and market instruments capable of automated parallel solid phase synthesis. As examples we can mention Nautilus, (http://www.argotech.com/products/other_products/nautilus.html), and Trident, (http://www.argotech.com/products/other_products/trident/trident.html) of Argonaut Technologies ***(92)***, Myriad of Mettler Toledo, (http://www.bohdan.com/mcs.htm) and SOPHAS of Zinsser Analytic ***(93)***, (http://zinsser-analytic.com). Even though these instruments are very sophisticated and can utilize a variety of reaction conditions, they are usually not particularly suitable for the automated repetitive multi-step synthesis of peptides (undeniably, peptides can be synthesized using these machines, but why should one use a Mercedes to go visit the neighbor, when a 5-min walk would do?). Most of these instruments are based on solid phase synthesis technology and use commercially available pipetting robots for the delivery of reagents and wash solutions to synthetic compartments. The capacity of these synthesizers ranges from 12 to 384 compounds that can be synthesized in one run. However, we will concentrate on the description of machines more or less designed for synthesis of peptides.

2.2.1. Synthesis Automation Based on Filtration

Single-vessel synthesizers were almost exclusively based on filtration. Much effort was spent on the design of an optimal reaction vessel that would accommodate various amounts of solid support. Vega (Tuscon, AZ; later Protein Technologies/Rainin), for example, adopted Merrifield's original design of a reaction vessel, and shaking by inverting the vessel, in its first synthesizer. That machine, however, performed only the washing and deprotection steps and all amino acid

Table 2
Automated Peptide Synthesizers

Company	Model	Capacity	Scale (mmol)	Chemistry	Comment
Advanced ChemTech[a]	ACT 384	384			
Advanced ChemTech	Apogee	10	0.1–0.5	Boc/Fmoc	Fast cycles
Advanced ChemTech	Velocity 16	16		Boc/Fmoc	Fast cycles, flow washing
Advanced ChemTech	ACT 90	2	0.05–35	Boc/Fmoc	
Advanced ChemTech	Apex 396	96	0.015–2	Boc/Fmoc	Overpressure filtration, vortex mixing
Advanced ChemTech	ACT 400	1	100–1000	Boc/Fmoc	
Advanced ChemTech	Vanguard	96		Boc/Fmoc	Only one step
Applied Biosystems[b]	ABI 433A	1	0.005–1	Boc/Fmoc	
Applied Biosystems	Pioneer	32	0.025–0.1	Fmoc	Continuous flow
Argonaut[c]	Quest 210	10	0.05–0.5	Boc/Fmoc	Only one step, manual
CEM[d]	Odyssey	1	0.005–5	Boc/Fmoc	Microwave, 12 consecutive
Chemspeed[e]	PSW1100	80	0.005–5	Boc/Fmoc	On-line cleavage and workup
CRB	PEPSYNthesizer	2	0.2–1	Fmoc	Not in production
CS Bio Co.[f]	CS336	3	0.05–0.25	Boc/Fmoc	Up to 108 AA
CS Bio Co.	CS736	1	2.0–25.0	Boc/Fmoc	
CS Bio Co.	CS936S	1	5–100	Boc/Fmoc	Customizable
CS Bio Co.	CS936	1	5–500	Boc/Fmoc	Customizable, mobile
Dan-Process[g]		1		Fmoc	Flow-through, industrial scale
Gilson	AMS422			Fmoc	Not in production
Heidolph Instruments[h]		24			Heating/cooling, single step only
Intavis AG[i]	ResPep	6	0.025–0.2	Fmoc	Column reactors
Intavis AG	ResPep Microscale	24	0.002–0.005	Fmoc	Column reactors
Intavis AG	MultiPep	192	0.002–0.01	Fmoc	Preactivation, filter multititer plates
Intavis AG	AutoSpot	800	0.000003	Fmoc	Only spotting automated

Intavis AG	MultiPep Spot	800	0.000003	Fmoc	Automated spot synthesis
Milligen/Biosearch	Milligen 9050	3	0.2–1	Fmoc	Not in production
Peptide Scientific[n]	PSI500	1	0.1–1000	Boc/Fmoc	Monitoring feedback
Pharmacia-LKB	Biolynx 4170	3	0.2–1	Fmoc	Not in production
Protein Technologies[j]	Symphony	12	0.005–0.350	Boc/Fmoc	12 independent protocols possible
Protein Technologies	Symphony-Cascade	12	0.05	Boc/Fmoc	Randomization chamber
Protein Technologies	Sonata	1	0.5-50	Boc/Fmoc	
Protein Technologies	PS3	1	0.01–0.25	Boc/Fmoc	Cartridges
Shimadzu	PSSM-8	96	0.005–0.4	Fmoc	Not in production
Spyder Instruments[k]	Compas 242	24	0.01–0.05	Fmoc	Centrifuge, tea-bags, not in production
Spyder Instruments	Compas 768	768	0.002–0.005	Fmoc	Centrifuge, 8 plates
Technikrom[l]		1		Fmoc	Flow-through, industrial scale
Zinsser Analytic[m]	SMPS350	144	0.05	Fmoc	First commercial multiple synthesizer
Zinsser Analytic	Pepsy-System	864	0.002	Fmoc	35 synthesis pens

[a] http://www.peptide.com
[b] http://www.appliedbiosystems.com
[c] http://www.argotech.com
[d] http://www.cem.com
[e] http://www.chemspeed.com
[f] http://www.csbio.com
[g] http://www.dan-process.dk/peptid.htm
[h] http://www.heidolph.com
[I] http://www.intavis.com
[j] http://pti-instruments.com
[k] http://www.5z.com/spyder
[l] http://www.technikrom.com
[m] http://zinsser-analytic.com
[n] http://www.peptidescientific.com

derivatives had to be added manually. The V-shape vessel of the Beckman model 990 synthesizer was capable of accepting various scales of synthesis and became a workhorse of several laboratories. Nowadays, most popular reaction vessels are cylindrical or spherical, usually made of glass with a volume range from several milliliters to tens of liters (CS Bio CS936 or Advanced ChemTech ACT400). Mixing in single-channel synthesizers is achieved by stirring, inert gas bubbling, vortexing, shaking, or liquid recirculation. Synthesizers based on the flow-through principle where solutions are recirculated through the stationary bed are quite rare; however, even this principle is used in an industrial-scale synthesizer (Technikrom, Dan-Process). Birr has built a synthesizer in which filtration was enabled by centrifugation of the vessel with porous walls ***(61)***.

Historically, the first commercially available automated multiple synthesizer, SMPS350, was brought to the market by Zinsser Analytic ***(70,94)*** in 1988. This machine was capable of the parallel synthesis of 144 peptides. Removal of the liquid from the solid phase slurry was performed by a needle equipped with stainless steel mesh that was consecutively immersed into each vessel. Obviously, this arrangement required washing of the needle after each immersion by backflushing with solvent and resulted in very long cycle times. The design of this synthesizer was "copied" by Advanced ChemTech and started a long array of synthesizers produced by this company.

One of the key issues in multiple solid phase synthesis is the parallel removal of excess reagent and wash solutions from the solid support in all synthetic compartments. In most currently available synthesizers this is achieved through the porous bottoms of the synthetic compartments, either by vacuum filtration (e.g., in instruments from Applied Biosystems), or application of pressure from the top of the compartments (e.g., in instruments from Advanced ChemTech). These methods bear the inherent risk of clogging of one or more compartments, resulting in insufficient liquid removal from the clogged compartments, overflow, and, consequently, contamination of neighboring compartments. This is especially dangerous in the case of enclosed systems such as the Teflon synthetic blocks of Advanced ChemTech instruments in which visual inspection of synthetic progress is not possible.

Chemspeed's PSW1100 offers synthesis in independent glass reactors (up to 80) of sizes from 13 mL to 100 mL. It is one of the most flexible synthesizers, allowing a variety of reaction conditions (preactivation, elevated temperature) and protocols and is capable of the automated cleavage of the final peptides, evaporation of the solution, and connecting to HPLC. **Figure 4** provides more details of this synthesizer.

Instead of building dedicated synthesizers, Tecan (http://www.tecan.com) decided (after the rather unsuccessful introduction of its own synthesizer) to

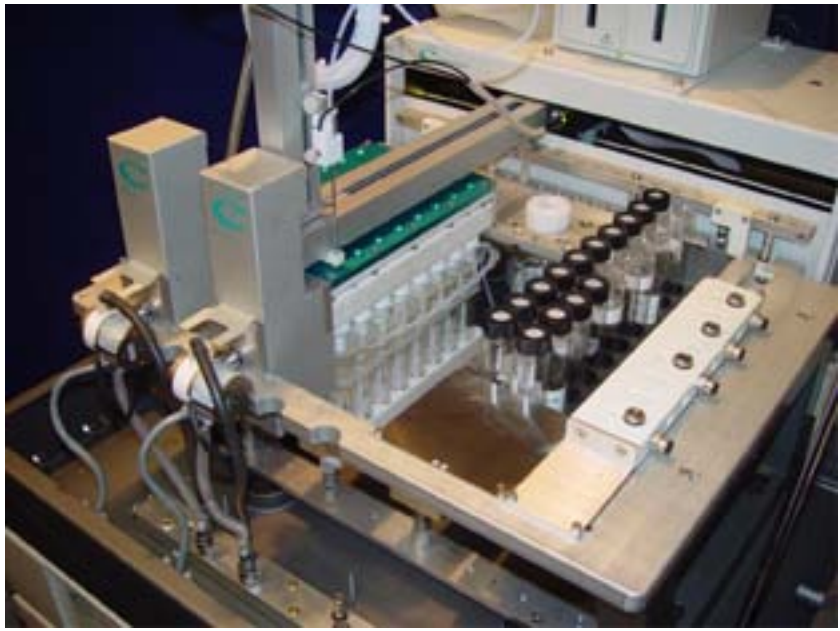


Fig. 4. Synthesis bed of Chemspeed's PSW1100. (Photo by ML.)

integrate the synthesis blocks of different producers (Robbins, Charybdis, H&P, Polyfiltronics, Bohdan, and others) with its own pipetting system. This flexible approach is definitely one to consider when building a dedicated peptide production facility.

The uniqueness of APLS1 ("Randomizer") is based on the fact that the 20 individual reaction compartments are connected into a larger continuous area. If there is a requirement to randomize (split and mix, or divide and recombine, in different author's terminology), the upper larger compartment is filled with the solvent, resin is pushed from the individual compartment into the common area by a flow of nitrogen, and the whole content is stirred. After sedimentation, the resin is uniformly distributed back into individual reaction chambers ***(95)***. Continuous stirring was used in the design of two synthesizers capable of resin randomization ***(85,96)***, one of which was commercialized but is now discontinued ***(82)***.

Zuckermann (Chiron) has developed a combinatorial synthesizer also capable of randomization ***(78,97,98)***. In this case, a Zymark robotic arm operated a syringe into which an aliquot of an "isopycnic" suspension of resin, after mixing all aliquots, was drawn and distributed into individual reaction compartments. One problem with this concept was fact that the swelling, and therefore

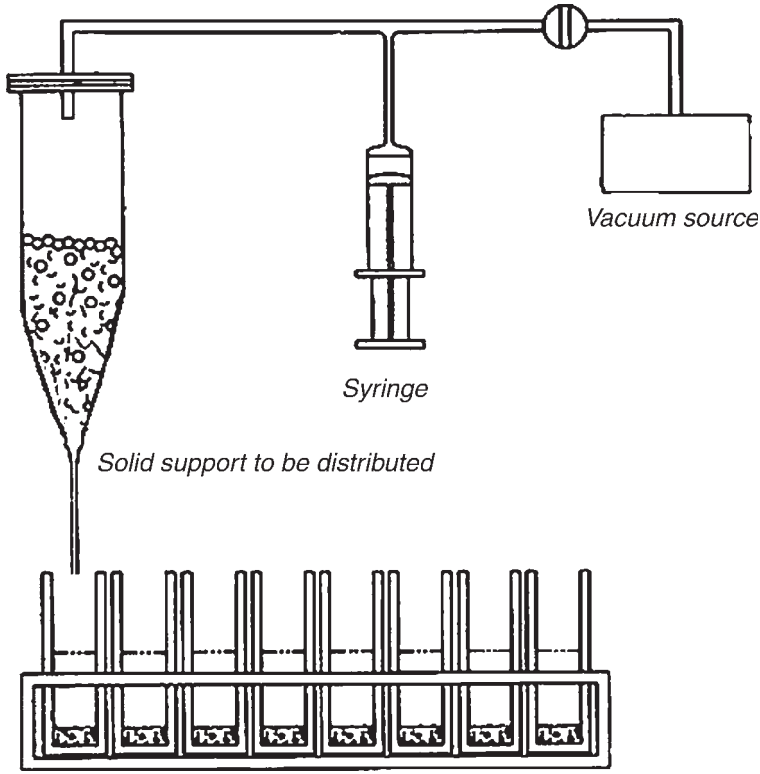


Fig. 5. Scheme of a tool for randomization by continuous bubbling.

the buoyancy, of the resin particles could change during the synthesis, the net result being uneven distribution of the aliquots.

Universal randomization and distribution of small volumes of resin particles can easily be achieved by the application of a continuously bubbled-through delivery vessel *(57)*. This principle of operation is illustrated in **Fig. 5**. A defined volume of the slurry is aspirated into a V-shape vessel with an opening at the bottom. Continuous suction through the vessel prevents slurry leakage and provides mixing. The gas connection has a T-junction to a syringe pump, which draws a gas volume equivalent to the volume to be dispensed. In the next step, suction through the vessel is stopped and the volume of gas in the syringe is immediately discharged, expelling an aliquot of slurry from the vessel. Suction is reapplied and the syringe is prepared for the delivery of the next aliquot.

The SPOT synthesizer mentioned earlier avoids the clogging problem because the filtration support becomes the synthetic substrate *(36)*. The size of the individual "spots" on which the synthesis is performed determines the scale and throughput of the synthesis. Usually, only nanomolar amounts of peptides are

Fig. 6. PepSy of Zinsser Analytic. Left—View of the system; Right—Tray of barcoded synthetic pens. (Photo by ML.)

generated in quantities up to 800 peptides per batch (1.5 mm spot size). Even though the peptides can be cleaved from the support and processed separately, very often the biological assay (e.g., binding) can be performed with peptides attached to the synthetic support ***(39,40)***.

The PepSy synthesizer, recently introduced to the market by Zinsser Analytic (http://www.zinsser-analytic.com), is shown in **Fig. 6**. In this synthesizer the solid support is sandwiched between two layers of porous material and reagents are delivered by the "synthetic pens." These pens are loaded with preactivated amino acid derivatives and the volume delivered to the support is defined by the time of contact of these pens with the porous cover of the synthetic chamber. The top of the pen is labeled with a two-dimensional bar code preventing misplacement of the reagent in the tray (positive recognition of the proper reagent by synthesizer is necessary before its use in synthesis). Even though this synthesizer is cleverly designed, it is not without problems. Only extremely small-scale synthesis can be performed because there is no means of mixing the support during the reaction, and the selection of the activated species is very limited—only preactivated derivatives can be used (e.g., fluorides). The recent modification of the synthesizer does not utilize resin embedded between first material, and therefore the pen is completely redesigned.

Synthesis in syringes also uses filtration as the separation principle. However, instead of total parallel processing, staggered partially parallel processing can also result in multiple synthesis. The synthesizer using this concept is probably closest to the concept of "robotic" synthesizers in the sense that it exactly mimics the process performed by the chemist handling the syringes manually ***(86)***. There are four types of syringes in this synthesizer: (1) syringes waiting to be included in the synthesis, (2) syringes in synthesis, (3) syringes with all steps of synthesis completed, and (4) syringes with reagents. A robotic arm

equipped with a specialized syringe gripper picks the syringe waiting to be included in the process, aspirates into it the appropriate reagent, shakes it, and places it on the tumbler for reaction. After that it checks whether any syringes in the process need attention and, if not, processes a new syringe. If the preparation of activated amino acid is required, the arm picks the syringe with the reagent, measures an appropriate volume of protected amino acid solution into the activation cup, adds the activator, and aspirates the freshly created mixture into a syringe ready for coupling. If the particular syringe has undergone all steps of synthesis, it is placed into the outgoing bin, where it awaits final processing. In this way, all syringes in the synthesis are in a different stage of processing, but the robotic arm is always busy, processing as many syringes as it can incorporate into its schedule. Long and short peptides can be synthesized at the same time, utilizing different protocols and scales.

Besides performing multiple syntheses in parallel, the accelerating of reaction rates may also achieve the same result—production of multiple peptides in the same time period. CEM (http://www.cem.com) has employed microwave irradiation to shorten both coupling and deprotection times. Their synthesizer, Odyssey (**Fig. 7**), achieves one cycle of peptide synthesis (Fmoc-based) in less than 10 min and is theoretically capable of the synthesis of 12 peptides in a row. However, for each new peptide the reaction vessel has to be cleaned and new batch of resin has to be transferred into it. This step raises concerns with the authors of this review about the possibility of cross-contamination.

An alternative method, avoiding filtration completely, employs aspiration of the liquid from the surface ***(42,99)***. This technology was automated and a robotic station was built that can process up to 72 microtiter plates (6912 compounds) in one batch ***(100)***. A robotic arm moves microtiter plates into stations into which the delivery of reagents is performed by 96-channel distributors and solvent aspiration is achieved by lifting the plate against an array of needles attached to a vacuum source.

2.2.2. Synthesis Automation Based on Centrifugation

Centrifugation is a powerful technique allowing the parallel processing of an unlimited number of reaction compartments ***(101)***. The first centrifugal multiple peptide synthesizer, Compas 242 ***(76,102)***, utilized centrifugation for liquid removal from the functionalized cotton used as the solid support or from resin contained in polypropylene mesh bags ***(103)***. This system enabled the automation of "tea-bag" synthetic methodology. In principle, however, separation of solid and liquid phases was still accomplished by filtration.

The key feature of the alternative centrifugation synthetic technology is a new method for the separation of the solid support from reagent solutions, termed "tilted plate centrifugation," which uses centrifugation as a means of liquid

Fig. 7. CEMs Odyssey microwave synthesizer. (Photo by ML.)

removal in conjunction with the use of tilted microtiter plates as reaction vessels. The tilted plate centrifugation technology greatly improves the earlier centrifugation method by using the wells of microtiter plates as synthetic compartments, thus enabling the parallel synthesis of much larger compound arrays (e.g., 3072 compounds when eight 384-well plates are used). The plates are mounted on a centrifugal plate and tilted slightly down toward the center of centrifugation, thus generating a pocket in each well, in which the solid support is collected during centrifugation, while the supernatant solutions are expelled from the wells (**Fig. 8**).

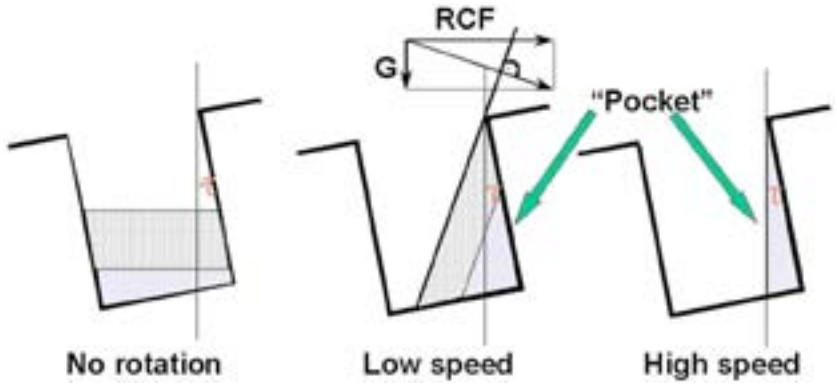


Fig. 8. Formation of the pocket in the well of a tilted plate during centrifugation (direction: left to right). The solid support (lower layer) is collected in the pocket, while the liquid (upper layer) is expelled from the well. The liquid surface angle is perpendicular to the resulting force vector of the relative centrifugal force (RCF) and gravity (G).

In order to ensure efficient liquid removal (i.e., no solution remaining in the wells after centrifugation), and at the same time avoid any loss of solid support during centrifugation, the volume of the well pockets should be equal to the volume of swollen resin in each well. This can be achieved by adjusting the pocket size by using plates with varying well volumes and/or modifying the tilt angle, as well as the speed of rotation.

An essential feature of this approach is that well-to-well cross-contamination with reagent solution or resin is avoided by the fact that the plates are tilted, while the direction of centrifugation is horizontal. Consequently, any liquid or resin expelled from the wells is either captured in the inter-well space of the plate or collected on the wall of the centrifugal drum. HPLC/MS analysis of all products prepared on the microtiter plate proved the fact that cross-contamination is not an issue.

Tilted-plate synthesis technology ***(26,101,104,105)*** is applicable not only to solid phase synthesis, but also to fluorous synthesis ***(106)***, solid and liquid-liquid extraction, and in-plate washing. A production facility (http://www.illumina.com/prod_oligos.asp) with a capacity of more than 30 million oligonucleotides (20 mers) per year was built on the basis of tilted centrifugation ***(101,107)***.

A high-throughput peptide synthesizer is shown in **Fig. 9**. Wash solutions and reagents common to all synthetic entities are automatically delivered through a multichannel distributor connected to a six-port selector valve. Building blocks and other specific reagents are individually delivered to their respective wells

Fig. 9. View of the deck of peptide synthesizer utilizing tilted centrifugation principle. (Photo by ML.)

by an integrated pipetting machine. Synthesis is performed in 96-well polypropylene microtiter plates. A critical factor for successful performance of the peptide synthesis is the guaranteed retention of all solid support in the washing cycles using the centrifugation process. The synthesis of a 20 mer peptide requires up to 300 centrifugation cycles, therefore, the loss of even a small fraction of the support in every step would result in the complete loss of the product. We have shown that even after several hundred centrifugations the solid support was retained.

During the evaluation of the high-throughput synthesizer we realized that it would be extremely beneficial to have an automated synthesizer mimicking the performance of the large machine but on a much lower scale. This allowed us to optimize the protocols and respond very quickly to the requests of researchers requiring only a small number of peptides. We found the optimal number of compounds synthesized in one batch to be 24 to 48. This number of synthetic compartments, wells, can be placed on the perimeter of a small rotor with the diameter of 14 cm. The rotor is then placed in the drum of the centrifuge covered by the lid with an integrated array of nozzles connected through the solenoid valves with vessels containing the particular reagents. During one rotation of

the rotor, all synthetic compartments can be placed under the appropriate nozzle and the reagent can be delivered. Individual amino acids can be delivered through one opening in the synthesizer deck by a pipetting machine. The removal of the solvent is achieved again by centrifugation. This "pet synthesizer" can produce a similar throughput compared to other expensive commercial instruments, but owing to the simplicity of its concept can be built for a fraction of the cost.

Centrifugation was also used for automation of one incarnation of SPOT synthesis. In this case the synthesis is performed on a functionalized polypropylene disk of the size of a conventional compact disk. The surface properties of this material allow for the discrete delivery and therefore synthesis at 2500 locations. After coupling, deprotection and washing are realized by delivering appropriate solvents close to the disk center and rapid rotation efficiently "sweeps" the liquid across the disk surface ***(108)***.

2.2.3. Other Multiple Synthesis Methods

For completeness, we should not forget about the photolithographic method pioneered by Fodor and co-workers ***(109)***. This technology utilizes selective deprotection of photocleavable amino protecting groups on the surface of a glass chip and exposure of the whole surface to the activated amino acid. Only deprotected locations accept the new amino acid and after removal of activated species, another set of locations can be deprotected and coupling with another amino acid can be performed. A disadvantage of this technique is the necessity to repeat the process for each synthetic step as many times as the number of amino acids that need to be coupled in a particular step—for natural peptides, 20 times. A major advantage is the extremely high density of peptides created on the glass surface (tens of thousands). Because of the complicated process of creating lithographic masks, this process is not easily adaptable in the average laboratory, and is exclusively used by Affymetrix (https://www.affymetrix.com) (even though described for peptide synthesis, the technology is now used almost exclusively for DNA synthesis). The masks were later replaced by the use of a digital micromirror projector for the deprotection of selected locations ***(110)***.

Alternative technology based on the same strategy (repetitive selective deprotection followed by global coupling) uses photochemically generated reagent for the *in situ* deprotection of amino groups of growing peptide attached to the chip surface ***(111)***. "Peptide chips" created by this technique are available from Xeotron (http://www.xeotron.com).

Synthesis of peptides in solution is currently practiced in very few laboratories, or is applied in the case of large-scale peptide synthesis. Multiple solution

synthesis is even more rare. An automated system for solution phase synthesis was described by Japanese authors and applied to the synthesis of only short peptides ***(112–114)***.

3. Conclusion

This chapter does not follow the usual format of this series; rather, it provides the reader with a general overview of techniques available for manual and automated high-throughput peptide synthesis. We were trying to be more specific in the description of techniques with which we are familiar, and that we believe are relevant for scientists in laboratories tasked to produce large numbers of peptides. In the case of automated synthesizers, we attempted to point out the potential problems that the user of a particular machine may face after the eventual purchase of the instrument—and that will definitely not be mentioned by the sales agent.

References

1. Lander, E. S., Linton, L. M., Birren, B., et al. (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860.
2. Venter, J. C., Adams, M. D., Myers, E. W., et al. (2001) The sequence of the human genome. *Science* **291,** 1304.
3. Aebersold, R. (2003) Constellations in a cellular universe. *Nature* **422,** 115–117.
4. Turecek, F. (2002) Mass spectrometry in coupling with affinity capture-release and isotope-coded affinity tags for quantitative protein analysis. *J. Mass Spectrom.* **37,** 1–14.
5. Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Amer. Chem. Soc.* **85,** 2149–2154.
6. Merrifield, R. B. (1985) Solid phase synthesis (Nobel lecture). *Angew. Chem. Int. Ed.* **24,** 799–810.
7. Merrifield, R. B. (1986) Solid phase peptide synthesis. *Science* **232,** 341–347.
8. Merrifield, R. B. (1993) *Life During a Golden Age of Peptide Chemistry: The Concept and Development of Solid-Phase Peptide Synthesis*. American Chemical Society, Washington, DC, pp. 1–297.
9. Merrifield, B. (1997) Concept and early development of solid-phase peptide synthesis. *Methods Enzymol.* **289,** 3–13.
10. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., and Knapp, R. J. (1991) A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **354,** 82–84.
11. Lam, K. S., Lebl, M., and Krchnak, V. (1997) The "one-bead one-compound" combinatorial library method. *Chem. Rev.* **97,** 411–448.
12. Valerio, R. M., Bray, A. M., Campbell, R. A., et al. (1993) Multipin peptide synthesis at the micromole scale using 2-hydroxyethyl methacrylate grafted polyethylene supports. *Int. J. Peptide Prot. Res.* **42,** 1–9.

13. Geysen, H. M., Meloen, R. H., and Barteling, S. J. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. USA* **81,** 3998–4002.
14. Bray, A. M., Maeji, N. J., and Geysen, H. M. (1990) The simultaneous multiple production of solution phase peptides; assesment of the Geysen method of simultaneous peptide synthesis. *Tetrahedron Lett.* **31,** 5811–5814.
15. Maeji, N. J., Valerio, R. M., Bray, A. M., Campbell, R. A., and Geysen, H. M. (1994) Grafted supports used with the multipin method of peptide synthesis. *React. Polym.* **22,** 203–212.
16. Carter, J. M., VanAlbert, S., Lee, J., Lyon, J., and Deal, C. (1992) Shedding light on peptide synthesis. *Biotechnology* **10,** 509–513.
17. Rasoul, F., Ercole, F., Pham, Y., et al. (2000) Grafted supports in solid-phase synthesis. *Biopolymers (Pept. Sci.)* **55,** 207–216.
18. Houghten, R. A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: Specificity of antigen-antibody interaction at the level of individual amino acids. *Proc. Natl. Acad. Sci. USA* **82,** 5131–5135.
19. Houghten, R. A., DeGraw, S. T., Bray, M. K., Hoffmann, S. R., and Frizzell, N. D. (1986) Simultaneous multiple peptide synthesis: The rapid preparation of large numbers of discrete peptides for biological, immunological, and methodological studies. *BioTechniques* **4,** 522–528.
20. Xiao, X., Zhao, C., Potash, H., and Nova, M. P. (1997) Combinatorial chemistry with laser optical encoding. *Angew. Chem. Int. Ed.* **36,** 780–782.
21. Nicolaou, K. C., Xiao, X. Y., Parandoosh, Z., Senyei, A., and Nova, M. P. (1995) Radiofrequency encoded combinatorial chemistry. *Angew. Chem. Int. Ed.* **34,** 2289–2291.
22. Moran, E. J., Sarshar, S., Cargill, J. F., et al. (1995) Radio frequency tag encoded combinatorial library method for the discovery of tripeptide-substituted cinnamic acid inhibitors of the protein tyrosine phosphatase PTP1B. *J. Amer. Chem. Soc.* **117,** 10787–10788.
23. Houghten, R. A., Bray, M. K., DeGraw, S. T., and Kirby, C. J. (1986) Simplified procedure for carrying out simultaneous hydrogen fluoride cleavages of protected peptide resins. *Int. J. Peptide Prot. Res.* **27,** 673–678.
24. Kerschen, A., Kanizsai, A., Botros, I., and Krchnak, V. (1999) Apparatus and method for cleavage of compounds from solid support by gaseous reagents. *J. Comb. Chem.* **1,** 480–484.
25. Lebl, M., Pires, J., Poncar, P., and Pokorny, V. (1999) Evaluation of gaseous hydrogen fluoride as a convenient reagent for parallel cleavage from the solid support. *J. Comb. Chem.* **1,** 474–479.
26. Lebl, M., Krchnak, V., Ibrahim, G., et al. (1999) Solid-phase synthesis of large tetrahydroisoquinolinone arrays by two different approaches. *Synthesis-Stuttgart* 1971–1978.
27. Blankemeyer-Menge, B. and Frank, R. (1988) Simultaneous multiple synthesis of protected peptide fragments on 'allyl'-functionalized cellulose disc supports. *Tetrahedron Lett.* **29,** 5871–5874.

28. Frank, R. and Doring, R. (1988) Simultaneous multiple peptide synthesis under continuous flow conditions on cellulose paper discs as segmental solid supports. *Tetrahedron* **44,** 6031–6040.
29. Frank, R., Heikens, W., Heisterberg-Moutsis, G., and Blocker, H. (1983) A new general approach for the simultaneous chemical synthesis of large numbers of oligonucleotides: Segmental solid supports. *Nucl. Acid. Res.* **11,** 4365–4377.
30. Dittrich, F., Tegge, W., and Frank, R. (1998) "Cut and combine": An easy membrane-supported combinatorial synthesis technique. *Bioorg. Med. Chem. Lett.* **8,** 2351–2356.
31. Eichler, J., Bienert, M., Stierandova, A., and Lebl, M. (1991) Evaluation of cotton as a carrier for solid phase peptide synthesis. *Peptide Res.* **4,** 296–307.
32. Jezek, J., Rinnova, M., and Lebl, M. (1993) Simultaneous multiple peptide synthesis: Comparison of T-bags and cotton. In *Peptides 1992*, Proc.22.EPS (Schneider, C. H. and Eberle, A. N., eds.), ESCOM, Leiden, pp. 306-307.
33. Lebl, M. and Eichler, J. (1989) Simulation of continuous solid phase synthesis: Synthesis of methionine enkephalin and its analogs. *Peptide Res.* **2,** 297–300.
34. Lebl, M. (1998) Solid-phase synthesis on planar supports. *Biopolymers (Pept. Sci.)* **47,** 397–404.
35. Stankova, M., Wade, S., Lam, K. S., and Lebl, M. (1994) Synthesis of combinatorial libraries with only one representation of each structure. *Peptide Res.* **7,** 292–298.
36. Frank, R. (1992) SPOT synthesis: An easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* **48,** 9217–9232.
37. Frank, R. and Overwin, H. (1996) SPOT synthesis. Epitope analysis with arrays of synthetic peptides prepared on cellulose membranes. *Methods Mol. Biol.* **66,** 149–169.
38. Frank, R., Hoffmann, S., Kiess, M., et al. (1996) Combinatorial synthesis on membrane supports by the SPOT technique: Imaging peptide sequence and shape space. In *Combinatorial Peptide and Nonpeptide Libraries: A Handbook* (Jung, G., ed.), VCH, Weinheim, Germany, pp. 363–386.
39. Wenschuh, H., Volkmer-Engert, R., Schmidt, M., Schulz, M., Schneider-Mergener, J., and Reineke, U. (2000) Coherent membrane supports for parallel microsynthesis and screening of bioactive peptides. *Biopolymers (Pept. Sci.)* **55,** 188–206.
40. Koch, J. and Mahler, M., Eds. (2002) *Peptide Arrays on Membrane Supports.* Springer, Berlin.
41. Eichler, J., Houghten, R. A., and Lebl, M. (1996) Inclusion volume solid-phase peptide synthesis. *J. Peptide Sci.* **2,** 240–244.
42. Krchnak, V., Weichsel, A. S., Lebl, M., and Felder, S. (1997) Automated solid-phase organic synthesis in micro-plate wells. Synthesis of N-(alkoxy-acyl)amino alcohols. *Bioorg. Med. Chem. Lett.* **7,** 1013–1016.
43. Wolfe, H. R. and Wilk, R. R. (1989) The RaMPS system: Simplified peptide synthesis for life science researchers. *Peptide Res.* **2,** 352–356.

44. Krchnak, V., Vagner, J., Flegel, M., and Mach, O. (1987) Continuous-flow solid-phase peptide synthesis. *Tetrahedron Lett.* **28,** 4469–4472.
45. Krchnak, V., Vagner, J., and Mach, O. (1989) Multiple continuous-flow solid-phase peptide synthesis. Synthesis of an HIV antigenic peptide and its omission analogues. *Int. J. Peptide Prot. Res.* **33,** 209–213.
46. Krchnak, V. and Vagner, J. (1990) Color-monitored solid-phase multiple peptide synthesis under low-pressure continuous flow conditions. *Peptide Res.* **3,** 182–193.
47. Vagner, J., Kocna, P., and Krchnak, V. (1991) Continuous-flow synthesis of agliadin peptides in an ultrasonic field and assay of their inhibition of intestinal sucrase activity. *Peptide Res.* **4,** 284–288.
48. Krchnak, V. and Vagner, J. (1992) Prediction and handling of difficult sequences in solid-phase peptide synthesis. In *Innovation and Perspectives in Solid Phase Synthesis.* (Epton, R., ed.), Intercept, Andover, UK, pp. 414–415.
49. Lebl, M. and Krchnak, V. (1997) Synthetic peptide libraries. *Methods Enzymol.* **289,** 336–392.
50. Baru, M. B., Cherskii, V. V., Danilov, A. V., Moshnikov, S. A., and Mustaeva, L. G. (1995) Automatic SynChrom system for solid phase peptide synthesis and liquid column chromatography. II. Application to solid phase peptide synthesis and liquid column chromatography. *Russ. J. Bioorch. Chem.* **21,** 506–516.
51. Baru, M. B., Cherskii, V. V., Danilov, A. V., Moshnikov, S. A., and Mustaeva, L. G. (1995) Automatic SynChrom system for solid phase peptide synthesis and liquid column chromatography. I. Principles of design and structural constituents. *Russ. J. Bioorch. Chem.* **21,** 498–505.
52. Baru, M. B., Mustaeva, L. G., Vagenina, I. V., Gorbunova, E. Y., and Cherskii, V. V. (2001) Pressure monitoring of continuous-flow solid-phase peptide synthesis. *J. Pept. Res.* **57,** 193–202.
53. Rodionov, I. L., Baru, M. B., and Ivanov, V. T. (1992) A swellographic approach to monitoring continuous-flow solid- phase peptide synthesis. *Peptide Res.* **5,** 119–125.
54. Krchnak, V., Vagner, J., Safar, P., and Lebl, M. (1988) Noninvasive continuous monitoring of solid phase peptide synthesis by acid-base indicator. *Collect. Czech. Chem. Commun.* **53,** 2542–2548.
55. Krchnak, V. and Padera, V. (1998) The domino blocks: A simple solution for parallel solid-phase organic synthesis. *Bioorg. Med. Chem. Lett.* **8,** 3261–3264.
56. Mjalli, A. M. M. and Toyonaga, B. E. (1995) Solid support combinatorial chemistry in lead discovery and SAR optimization; http://www.netsci.org/Science/Combichem/feature03.html. *Net. Sci.* **1**.
57. Lebl, M., Pokorny, V., and Krchnak, V. (2000) Apparatus and method for combinatorial chemistry synthesis. Trega Biosciences, Inc. San Diego, CA. US Patent 6,045,755. Lebl, M. and Krchnak, V. (2004) *J. Comb. Chem.* (in press).
58. Merrifield, R. B., Stewart, J. M., and Jernberg, N. (1966) Instrument for automated synthesis of peptides. *Anal. Chem.* **38,** 1905–1914.

59. Merrifield, R. B. and Stewart, J. M. (1965) Automated peptide synthesis. *Nature* **207,** 522–523.
60. Brunfeldt, K. (1973) Automation in solid phase peptide synthesis. In *Peptides 1972,* Proc.12.EPS (Hanson, H. and Jakubke, H. D., eds.), North-Holland Publishing Company, Amsterdam, pp. 141–151.
61. Birr, C. (1978) Automatization of the Merrifield peptide synthesis. In *Aspects of the Merrifield Peptide Synthesis* (Birr, C., ed.), Springer-Verlag, Berlin; New York, pp. 72–80.
62. Edelstein, M. S., McNair, D. S., and Sparrow, J. T. (1981) The conversion of solid phase peptide synthesizers to computer control. In *Peptides: Synthesis, Structure, Function* (Rich, D. H. and Gross, E., eds.), Pierce Chemical Company, Rockford, IL, pp. 217–220.
63. Jonczyk, A. and Meienhofer, J. (1983) Automated flow reactor synthesizer for fast synthesis of peptides using Fmoc protection. In *Peptides: Structure and Function,* Proc.8.APS (Hruby, V. J. and Rich, D. H., eds.), Pierce Chemical Company, Rockford, IL, pp. 73–77.
64. Blaha, I., Zaoral, M., Krchnak, V., Jehnicka, J., Stepanek, J., and Kalousek, J. (1986) Automatic device for solid-phase peptide synthesis. *Chem. Listy* **80,** 994.
65. Cameron, L. R., Holder, J. L., Meldal, M., and Sheppard, R. C. (1988) Peptide synthesis. Part 13. Feedback control in solid phase synthesis. Use of fluorenylmethoxycarbonyl amino acid 3,4-dihydro- 4-oxo-1,2,3-benzotriazin-3-yl esters in a fully automated system. *J. Chem. Soc. Perkin Trans.* **1**, 2895–2901.
66. Geiser, T., Beilan, H., Bergot, B. J., and Otteson, K. M. (1988) Automation of solid-phase peptide synthesis. In *Macromolecular Sequencing and Synthesis: Selected Methods and Applications* (Schlesinger, D. H., ed.), Alan R. Liss, New York, pp. 199–218.
67. Newton, R., Fox, J. E., and Mizrahi, A. (1988) Automation of peptide synthesis. *Synt. Peptide. Biotechnol.* 1–24.
68. Bridgham, J., Geiser, T. G., Hunkapiller, M. W., et al. (1989). Automated polypeptide synthesis process. Applied Biosystems, Inc. Foster City, CA. US Patent 4,816,513.
69. Kearney, T. and Giles, J. (1989) Fmoc peptide synthesis with a continuous flow synthesizer. *Amer. Biotechnol. Lab.* **7,** 34–44.
70. Schnorrenberg, G. and Gerhardt, H. (1989) Fully automatic simultaneous multiple peptide synthesis in micromolar scale: Rapid synthesis of series of peptides for screening in biological assays. *Tetrahedron* **45,** 7759–7764.
71. Gausepohl, H., Kraft, M., Boulin, C., and Frank, R. W. (1990) A robotic workstation for automated multiple peptide synthesis. In *Innovations and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), SPCC, Birmingham, UK, pp. 487–490.
72. Judd, A. K. (1991) Multiple polymer synthesizer. SRI International, Menlo Park, CA. US Patent 5,053,454.
73. Schnorrenberg, G., Wiesmuller, K. H., Beck-Sickinger, A. G., Drechsel, H., and Jung, G. (1991) Rapid fully automatic SMPS for epitope mapping of influenza

nucleoprotein. In *Peptides 90,* Proc.21.EPS (Giralt, E. and Andreu, D., eds.), ESCOM, Leiden, pp. 202–203.
74. Fox, J. E. (1992) Automatic multiple peptide synthesis. *Biochem. Soc. Trans.* **20,** 851–853.
75. Gausepohl, H., Boulin, C., Kraft, M., and Frank, R. W. (1992) Automated multiple peptide synthesis. *Peptide Res.* **5,** 315–320.
76. Lebl, M., Stierandova, A., Eichler, J., et al. (1992) An automated multiple solid phase peptide synthesizer utilizing cotton as a carrier. In *Innovation and Perspectives in Solid Phase Peptide Synthesis* (Epton, R., ed.), Intercept Limited, Andover, UK, pp. 251–257.
77. Nokihara, K., Yamamoto, R., Hazama, M., Wakizawa, O., and Nakamura, S. (1992) Design and applications of a novel simultaneous multiple solid phase peptide synthesizer. In *Innovation and Perspectives in Solid Phase Peptide Synthesis* (Epton, R., ed.), Intercept Limited, Andover, UK, pp. 445–448.
78. Zuckermann, R. N., Siani, M. A., and Banville, S. C. (1992) Control of the zymate robot with an external computer: Construction of a multiple peptide synthesizer. *Lab. Robotics Automation* **4,** 183–192.
79. Bridgham, J., Geiser, T., Hunkapiller, M. W., et al. (1993). Automated polypeptide synthesis apparatus. Applied Biosystems, Inc. Foster City, CA. US Patent 5,186,898.
80. Neimark, J. and Briand, J. P. (1993) Development of a fully automated multichannel peptide synthesizer with integrated TFA cleavage capability. *Peptide Res.* **6,** 219–228.
81. Saneii, H. H., Shannon, J. D., Miceli, R. M., Fischer, H. D., and Smith, C. W. (1994) Fully automated selection and synthesis of peptide libraries. In *Peptides: Chemistry, Structure and Biology,* Proc.13.APS (Hodges, R. S. and Smith, J. A., eds.), ESCOM, Leiden, pp. 1018–1020.
82. Saneii, H. H. and Shannon, J. D. (1994) Fully automated solid phase synthesis of combinatorial libraries on the peptide librarian. In *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Intercept, Andover, UK, pp. 335–338.
83. Chang, H. W. and Slavazza, D. M. (1995) Solid phase peptide synthesizer. US Patent 5,453,487.
84. Nokihara, K., Hazama, M., Yamamoto, R., and Nakamura, S. (1995) Simultaneous multiple chemical synthesizer. Shimadzu Corporation, Kyoto Japan. US Patent 5,395,594.
85. Boutin, J. A. and Fauchere, J. L. (1996) Second-generation robotic synthesizer for peptide, pseudopeptide and non-peptide libraries. In *Proceedings of the International Symposium on Laboratory Automation and Robotics 1995* (Little, J. N., O'Neil, C., and Strimaitis, J. R., eds.), Zymark Corp., Hopkinton, MA, pp. 197–210.
86. Krchnak, V., Cabel, D., and Lebl, M. (1996) MARS: Multiple automated robotic synthesizer for continuous flow of peptides. *Peptide Res.* **9,** 45–49.
87. Daniels, S. B., Hantman, S. F., Sole, N. A., Gibney, B. R., Rabanal, F., and Kates, S. A. (1998) Pioneer(TM): A continuous-flow peptide synthesis system. In *Pep-*

tides 1996: Proceedings of the Twenty-Fourth European Peptide Symposium (Ramage, R. and Epton, R., eds.), Mayflower Scientific Ltd., Kingswinford, UK, pp. 323–324.
88. Carpino, L. A. and Han, G. Y. (1970) The 9-fluorenylmethoxycarbonyl function, a new base-sensitive aminoprotecting group. *J. Amer. Chem. Soc.* **92,** 5748–4749.
89. Atherton, E. and Sheppard, R. C. (1989) *Solid Phase Peptide Synthesis: A Practical Approach*, IRL Press at Oxford University Press, Oxford; New York, pp. 1–203.
90. Dryland, A. and Sheppard, R. C. (1986) Peptide synthesis. Part 8. A system for solid-phase synthesis under low pressure continuous flow conditions. *J. Chem. Soc. Perkin Trans.* **1,** 125–137.
91. Cargill, J. F. and Lebl, M. (1997) New methods in combinatorial chemistry: Robotics and parallel synthesis. *Curr. Opin. Chem. Biol.* **1,** 67–71.
92. Gooding, O., Hoeprich, P. D. Jr., Labadie, J. W., Porco, J. A. Jr., van Eikeren, P., and Wright, P. (1996) Boosting the productivity of medicinal chemistry through automation tools: Novel technological developments enable a wide range of automated synthetic procedures. In *Molecular Diversity and Combinatorial Chemistry. Libraries and Drug Discovery* (Chaiken, I. M. and Janda, K. D., eds.), American Chemical Society, Washington, DC, pp. 199–206.
93. Zinsser, W. (2000) SOPHAS—A real high throughput synthesizer. In *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Mayflower Worldwide, Kingswinford, UK, pp. 61–66.
94. Zinsser, W. (2002) Workbench automation in synthesis: From preparation to final substance. In *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Mayflower Worldwide, Kingswinford, UK, pp. 67–78.
95. Bartak, Z., Bolf, J., Kalousek, J., et al. (1994) Design and construction of the automatic peptide library synthesizer. *Methods: A Companion to Methods in Enzymology* **6,** 432–437.
96. Boutin, J. A., Hennig, P., Lambert, P. H., et al. (1996) Combinatorial peptide libraries: Robotic synthesis and analysis by nuclear magnetic resonance, mass spectrometry, tandem mass spectrometry, and high-performance capillary electrophoresis techniques. *Anal. Biochem.* **234,** 126–141.
97. Zuckermann, R. N., Kerr, J. M., Siani, M. A., and Banville, S. C. (1992) Design, construction and application of a fully automated equimolar peptide mixture synthesizer. *Int. J. Peptide Prot. Res.* **40,** 497–506.
98. Zuckermann, R. N. and Banville, S. C. (1992) Automated peptide-resin deprotection/cleavage by a robotic workstation. *Peptide Res.* **5,** 169–174.
99. Lebl, M., Pokorny, V., and Krchnak, V. (2000) Apparatus and method for combinatorial chemistry synthesis. Trega Biosciences, Inc. San Diego, CA. US Patent 6,045,755.
100. Lebl, M. and Krchnak, V. (1999) Techniques for massively parallel synthesis of small organic molecules. In *Innovation and Perspectives in Solid Phase Synthesis & Combinatorial Libraries* (Epton, R., ed.), Mayflower Scientific Limited, Birmingham, UK, pp. 43–46.

101. Lebl, M. (2003) Centrifugation based automated synthesis technologies. *J. Assoc. Lab. Autom.* **8,** 30–36.
102. Pokorny, V., Mudra, P., Jehnicka, J., et al. (1994) Compas 242. New type of multiple peptide synthesizer utilizing cotton and tea bag technology. In *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Mayflower Worldwide Limited, Birmingham, UK, pp. 643–648.
103. Bolf, J., Eichler, J., Jehnicka, J., et al. (1993) Multiple synthesis of peptide(s) on solid carrier. Ceskoslovenska Akademie Ved, UOCHB Prague CR. CS US Patent 5,202,418; 5,338,831; 5,342,585.
104. Lebl, M. (2000) Method for separation of liquid and solid phases for solid phase organic syntheses. Trega Biosciences, Inc. San Diego, CA. US Patent 6,121,054.
105. Lebl, M. (1999) New technique for high-throughput synthesis. *Bioorg. Med. Chem. Lett.* **9,** 1305–1310.
106. Studer, A. and Curran, D. P. (1997) A strategic alternative to solid phase synthesis: Preparation of a small isoxazoline library by "fluorous synthesis." *Tetrahedron* **53,** 6681–6696.
107. Lebl, M., Burger, C., Ellman, B., et al. (2001) Fully automated parallel oligonucleotide synthesizer. *Collect. Czech. Chem. Commun.* **66,** 1299–1314.
108. Adler, F., Turk, G., Frank, R., et al. (2000) A new array format for the automated parallel combinatorial synthesis by the SPOT-technique. In *Innovation and Perspectives in Solid Phase Synthesis* (Epton, R., ed.), Mayflower Worldwide, Kingswinford, UK, pp. 221–222.
109. Fodor, S. P. A., Leighton, R. J., Pirrung, M. C., Stryer, L., Lu, A. T., and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* **251,** 767–773.
110. Singh-Gasson, S., Green, R. D., Yue, Y., et al. (1999) Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat. Biotech.* **17,** 974–978.
111. Pellois, J. P., Wang, W., and Gao, X. (2000) Peptide synthesis based on t-Boc chemistry and solution photogenerated acids. *J. Comb. Chem.* **2,** 355–360.
112. Kuroda, N., Hattori, T., Fujioka, Y., Cork, D. G., Kitada, C., and Sugawara, T. (2001) Application of automated synthesis suite to parallel solution-phase peptide synthesis. *Chem. Pharm. Bull. Tokyo* **49,** 1147–1154.
113. Sugawara, T., Kobayashi, K., Okamoto, S., Kitada, S., and Fujino, M. (1995) Application of unique automated synthesis system for solution-phase peptide synthesis. *Chem. Pharm. Bull. Tokyo* **43,** 1272–1280.
114. Kuroda, N., Hattori, T., Kitada, C., and Sugawara, T. (2001) Solution-phase automated synthesis of tripeptide derivatives. *Chem. Pharm. Bull. Tokyo* **49,** 1138–1146.

12

Backbone Amide Linker Strategies for the Solid-Phase Synthesis of C-Terminal Modified Peptides

Jordi Alsina, Steven A. Kates, George Barany, and Fernando Albericio

Summary

This chapter describes backbone amide linker (BAL) strategies for the N^{α}-Fmoc solid-phase synthesis of C-terminal modified peptides. Most solid-phase protocols for the assembly of such peptides have limited generality, because they rely on the C^{α}-carboxyl for attachment to the solid support. In the BAL approach, the growing peptide chain is anchored through a backbone nitrogen, thus allowing significant flexibility for chemical modification of the C-termini. In effect, any peptide containing C-terminal variations can be prepared in overall good purity and yield, with minimal side reactions, by using one or more of three variations (original and two modifications) of the BAL strategy.

Key Words: Acidolytic cleavage; backbone amide linkage; bioconjugate; chemical ligation; combinatorial chemistry; handle; linker; peptide alcohol; peptide aldehyde; peptide ester; peptide thioester; protecting group; solid support.

1. Introduction

Solid-phase synthesis (SPS) of peptides containing at their C-termini the usual carboxylic acid or carboxamide functionalities is a well-established process; the peptide is traditionally attached to the resin through the α-carboxyl group of the C-terminal residue, and synthesis proceeds in the $C \rightarrow N$ direction ***(1,2)***. However, synthetic peptides containing modifications at the C-termini are often desired because of their potential therapeutic properties and/or synthetic significance as intermediates in peptide and protein chemistry. Therefore, effective solid-phase methods are needed for the preparation of these peptide targets ***(3)***. The present chapter describes backbone amide linker (BAL) strategies ***(4)***

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

for SPS of several classes of C-terminal-modified peptides, all based on the concept that the growing peptide is attached to an appropriate solid support through a backbone amide nitrogen. Since the C-terminal carboxyl group is not involved in the anchoring of the peptide to the support, the BAL strategies allow considerable flexibility to manipulate the C-terminal functionality while the peptide remains attached to the solid support.

The original BAL strategy ***(5)*** (**Fig. 1**) was implemented with the tris(alkoxy)-benzylamide (PAL) system ***(6)***. Thus, the backbone amide linkage is established by a reductive amination reaction involving the α-amine of the intended C-terminal residue (or of an appropriately protected C-terminal modified derivative) and an aldehyde-functionalized solid support (e.g., 5-(4-formyl-3,5-dimethoxyphenoxy)valeramide resin ≡ BAL-resin), followed by *N*-acylation under specialized conditions using the N^{α}-protected second amino acid residue. Further peptide elongation continues using N^{α}-9-fluorenylmethyloxycarbonyl (Fmoc) amino acid building blocks, and final acidolytic cleavage with trifluoroacetic acid (TFA) releases the desired free (unprotected) peptide into solution. Several C-terminal-modified amino acid derivatives, with C-terminal protecting and/or modifying groups entirely stable to conditions for repeated Fmoc removal, have been used, such as aliphatic esters ***(5)***, *tert*-butyl ethers ***(5)***, dimethyl acetals ***(5)*** and 1,3-dioxolanes ***(7)***, *N*,*N*-dialkylamides ***(5)***, and trithioortho esters ***(8)*** for the synthesis of peptide esters, alcohols, aldehydes, *N*,*N*-dialkylamides, and thioesters, respectively.

Two extensions of BAL methodology have been described that allow preparation of C-terminal-modified peptides containing hindered, unreactive, and/or base-sensitive moieties ***(9)***, or circumvent other synthetic challenges ***(10)***.

As a prototype of the first modified BAL strategy (**Fig. 2**), peptides containing relatively non-nucleophilic and/or piperidine-labile C-terminal modifications (specifically, *p*-nitroanilides and thioesters) were prepared by incorporating the desired C-terminal modified residue at the end of the synthesis, immediately prior to the final cleavage/deprotection step. To do this, the synthesis starts by anchoring the penultimate residue, with the corresponding C^{α}-carboxyl group orthogonally protected as an allyl ester, through the backbone nitrogen. Peptide assembly in the $C \rightarrow N$ direction using standard Fmoc procedures is followed by selective removal of the C-terminal allyl ester protecting group, activation of the pendant carboxyl thus released, coupling with the appropriate C-terminal modified amino acid derivative, and final cleavage with TFA.

The second modified BAL strategy accommodates essentially any C-terminal modification stable to piperidine, and is illustrated here (**Fig. 3**) for the preparation of peptide *N*,*N*-dialkylamides containing prolyl or *N*-alkylamino acyl derivatives at the C-terminus. (Use of the original BAL approach for the synthesis of these target peptides is inappropriate because the initial reductive

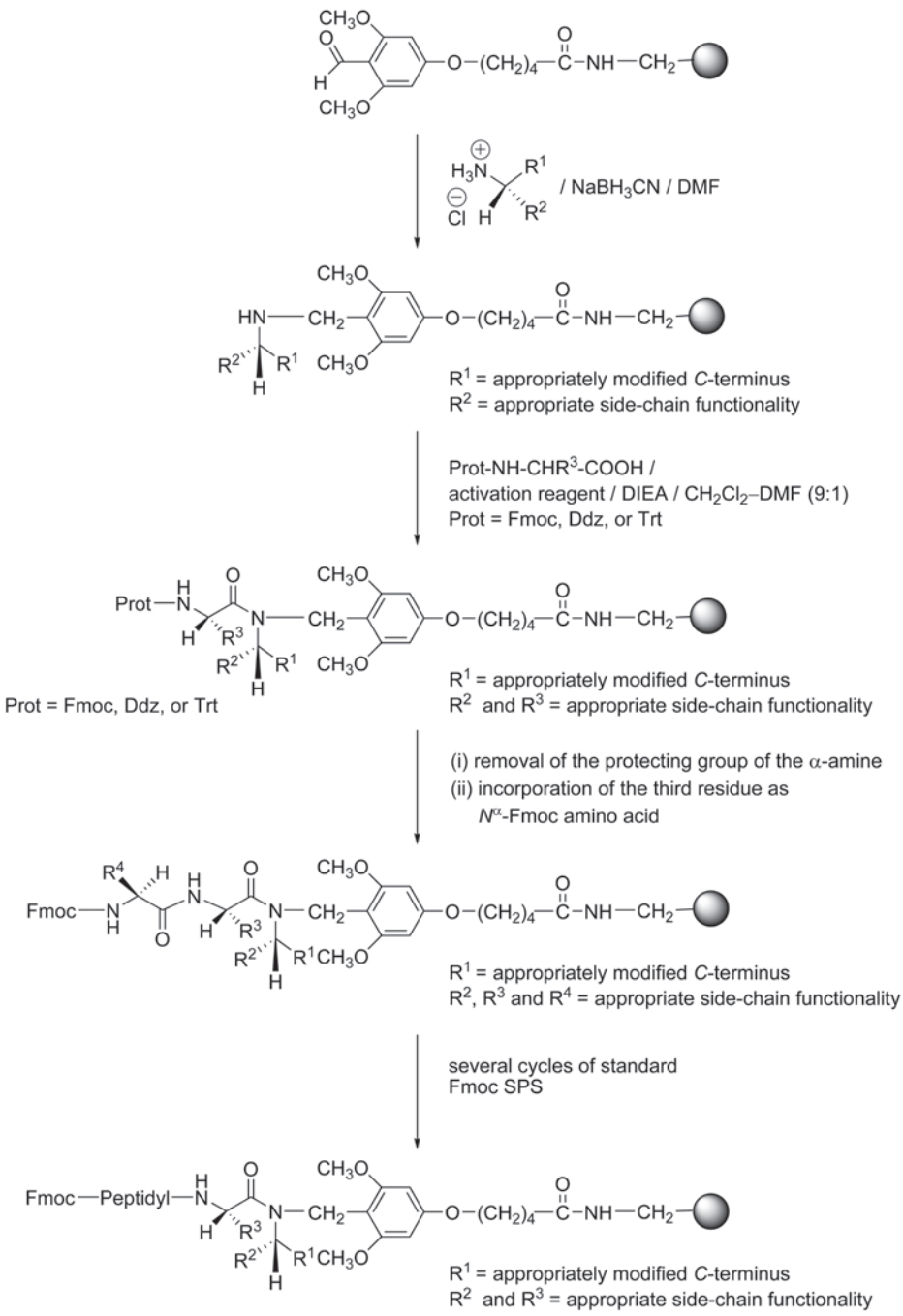


Fig. 1. Synthesis of resin-bound C-terminal modified peptides using original BAL strategy.

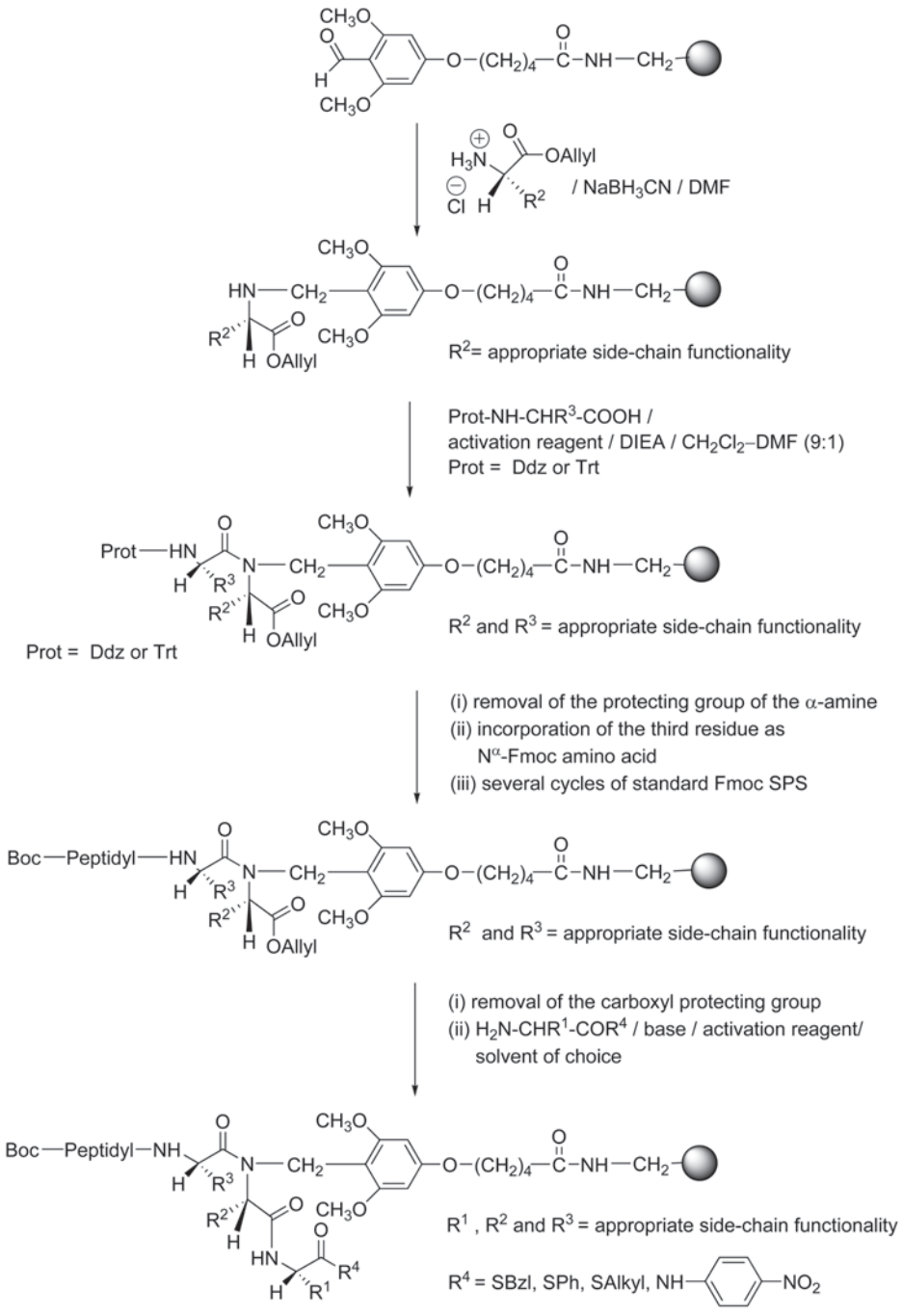


Fig. 2. Synthesis of resin-bound C-terminal modified peptides using BAL strategy modification I.

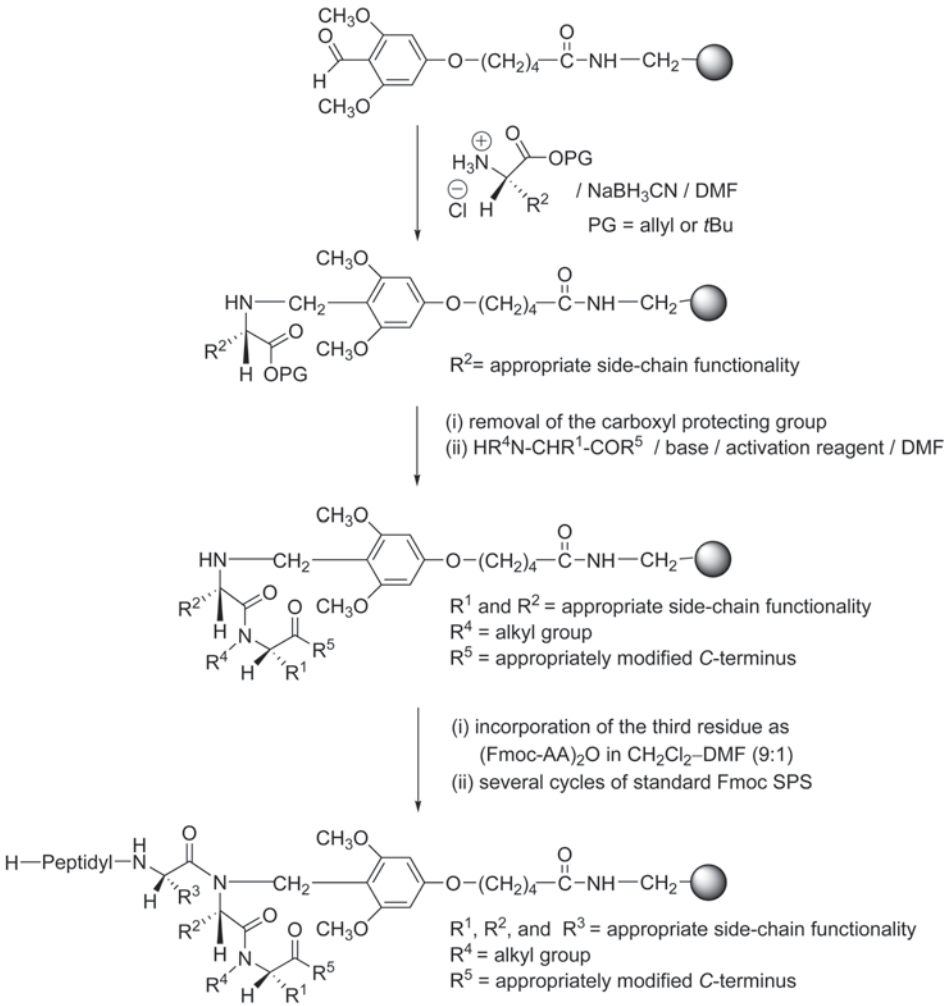


Fig. 3. Synthesis of resin-bound C-terminal modified peptides using BAL strategy modification II.

amination step would involve a secondary amine; the first modified BAL strategy is not applicable either, because of the risk of racemization at the penultimate residue after it is activated and coupled relatively slowly with incoming *N*-alkylamino acids.) In the first step, BAL anchoring is applied to the eventual penultimate residue of the target peptide, with its C^{α}-carboxyl group protected as an allyl or *tert*-butyl ester. Selective removal of the C^{α}-carboxyl protecting group—without premature cleavage of the BAL anchor—is followed by activation of the carboxyl using a phosphonium salt and coupling with the desired C-

terminal residue (racemization is essentially avoided). In this strategy, the two key steps just mentioned involve manipulation of the resin-bound (penultimate) residue as the free amine, as opposed to the amide as in previous strategies. The final steps are, as before, acylation of the BAL-anchored amine, peptide chain elongation by Fmoc protocols, and TFA cleavage.

2. Materials

1. BAL-resin ≡ 5-(4-formyl-3,5-dimethoxyphenoxy)valeramide resin (Applied Biosystems, Framingham, MA).
2. Polyethylene glycol-polystyrene graft resin (PEG-PS) (Applied Biosystems).
3. 4-Methylbenzhydrylamine polystyrene resin (MBHA-PS) (Calbiochem-Novabiochem AG, San Diego, CA).
4. Fmoc-amino acids, commercially available from different sources (e.g., Novabiochem; Advanced ChemTech, Louisville, KY; Bachem Bioscience, King of Prussia, PA).
5. Amino acid *tert*-butyl ester hydrochloride salts (Advanced ChemTech).
6. Amino acid *p*-nitroanilide hydrochloride salts and *N,N*-dimethylamide hydrochloride salts (Bachem Bioscience).
7. Ddz-amino acids (Advanced ChemTech).
8. HATU ≡ 1-[bis(dimethylamino)methylene]-1H 1,2,3-triazolo-[4,5,6]pyridinium hexafluorophosphate 3-oxide (Applied Biosystems).
9. PyAOP ≡ 7-azabenzotriazol-1-yl-*N*-oxytris(pyrrolidino)phosphonium hexafluorophosphate (Applied Biosystems).
10. PyBOP ≡ benzotriazol-1-yl-*N*-oxytris(pyrrolidino)phosphonium hexafluorophosphate (Calbiochem-Novabiochem).
11. Solvents (e.g., dichloromethane [CH_2Cl_2]; *N,N*-dimethylformamide [DMF]) and other reagents (acetic acid [HOAc]; *N,N′*-diisopropylcarbodiimide [DIPCDI]; *N,N*-diisopropylethylamine [DIEA]; 1-hydroxybenzotriazole [HOBt]; *N*-methylmorpholine [NMM]; piperidine; sodium cyanoborohydride [$NaBH_3CN$]; sodium *N,N*-diethyldithiocarbamate; tetrakistriphenylphosphine palladium (0) [$Pd(PPh_3)_4$]; triethylsilane [Et_3SiH]; trifluoroacetic acid [TFA]) are commercially available from different sources (e.g., Aldrich Chemical Co, Milwaukee, WI; and Fisher Scientific, Pittsburgh, PA).

3. Methods

This section outlines experimental procedures for the three BAL strategies.

3.1. Original BAL Strategy

The required steps for SPS of C-terminal-modified peptides using the originally envisaged BAL strategy (**Fig. 1**) are described in **Subheadings 3.1.1.–3.1.5.** These include (1) reductive amination procedure for the incorporation of appropriately protected C-terminal-modified amino acid derivatives into BAL; (2) acylation of the secondary amine-BAL-resin intermediates with an

N^{α}-protected penultimate amino acid; (3) protocol for introduction of the second and third residues when the sequence poses a risk of serious chain loss due to diketopiperazine (DKP) formation; (4) stepwise chain elongation using Fmoc protocols to complete on-resin assembly of the desired C-terminal-modified peptide; and (5) acidolytic cleavage to release the peptide into solution.

3.1.1. Reductive Amination Procedure

An optimized procedure for reductive amination involves mixing BAL-resin (1 eq) (*see* **Note 1**) with excess (10 eq) of the hydrochloride salt of the amino acid derivative plus (10 eq) sodium cyanoborohydride, simultaneously and without pre-equilibration (*see* **Note 2**) in DMF. Under these neutral or slightly acidic conditions, the reaction proceeds to completion, without any detectable racemization. A detailed experimental procedure for a specific example, taken from **ref. *5***, follows: BAL-Ile-4-methylbenzhydrylamine polystyrene resin (MBHA-PS) (*see* **Note 3**) (50 mg, 0.48 mmol/g, 1 eq) (*see* **Note 4**) was washed with CH_2Cl_2 (2×0.5 min, 3 mL each) and DMF (2×0.5 min, 3 mL each). HCl•H-Leu-O*t*Bu (54 mg, 10 eq) and 15 mg (10 eq) $NaBH_3CN$ were dissolved in 0.4 mL DMF, combined, added to the resin, and the reaction was allowed to proceed at 25°C for 1 h. The resin was then washed with DMF (5×0.5 min, 3 mL each), CH_2Cl_2 (3×0.5 min, 3 mL each), DMF (3×0.5 min, 3 mL each), piperidine–DMF (1:4, 3×1 min, 3 mL each), DMF (5×0.5 min, 3 mL each), and CH_2Cl_2 (3×0.5 min, 3 mL each), and dried.

3.1.2. Acylation of Secondary Amine-BAL-Resin Intermediates

The next reaction required for stepwise synthesis, *N*-acylation of the resulting secondary α-amine attached to BAL-resin by an appropriately protected second amino acid residue, is more challenging than comparable reactions of unsubstituted primary amines. Near-quantitative yields for acylation of the sterically hindered amines are obtained by using either N^{α}-protected amino acids as preformed symmetrical anhydrides in CH_2Cl_2–DMF (9:1), or N^{α}-protected amino acids activated by HATU in the presence of DIEA in CH_2Cl_2–DMF (9:1). Activation reagents other than HATU that give satisfactory results, always using CH_2Cl_2–DMF (9:1) as solvent to assure good coupling yield (*see* **Note 5**), are 1,1,3,3-tetramethyl-2-fluoroformamidinium hexafluorophosphate (TFFH), PyAOP, and bromotris(pyrrolidino)phosphonium hexafluorophosphate (PyBroP). Detailed experimental procedures for two specific examples, again from ***(5)***, follow:

1. Using preformed symmetrical anhydrides: 83 mg (10 eq) Fmoc-Phe-OH was dissolved in 0.5 mL CH_2Cl_2–DMF (9:1), and 20 µL (6 eq) DIPCDI was added. After a 5-min preactivation period, the anhydride solution was filtered and added to 50 mg (0.43 mmol/g, 1 eq) (H-BAL-Ile-MBHA-PS)-Leu-O*t*Bu, and coupling was allowed

to proceed for 2 h. The resin was washed subsequently with CH_2Cl_2 (5 × 0.5 min, 3 mL each), then DMF (5 × 0.5 min, 3 mL each), and the coupling procedure was repeated (2 h).

2. Using HATU as an activation reagent: 83 mg (10 eq) Fmoc-Phe-OH was dissolved in 0.5 mL CH_2Cl_2–DMF (9:1), followed by 73 μL (20 eq) DIEA, and then this solution was added to 50 mg (0.43 mmol/g, 1 eq) H-(BAL-Ile-MBHA-PS)-Leu-O*t*Bu. Following 30 s of stirring, HATU 81 mg (10 eq) was added as a solid to the peptidyl-resin solution. The coupling was allowed to proceed for 2 h. The resin was washed with CH_2Cl_2 (5 × 0.5 min, 3 mL each), DMF (5 × 0.5 min, 3 mL each), and the coupling procedure was repeated (2 h).

3.1.3. Introduction of Second and Third Residues When Sequence Poses Risk of Serious Chain Loss Owing to DKP Formation

When an amino acid *n*-alkyl or allyl ester is used as the first building block, after the Fmoc group is removed at the dipeptidyl level, significant DKP formation can occur. This side reaction can be minimized by using (instead of Fmoc) an acidolytically removable N^{α}-amino protecting group (either 2-(3,5-dimethoxyphenyl)propyl(2)oxycarboxyl [Ddz] or trityl [Trt]) for the second residue. In this way, the amine end group of the BAL-anchored dipeptide remains protonated after deprotection, and is therefore unreactive until the next step, a coupling with *in situ* neutralization ***(11)***. A detailed experimental procedure for a specific example, again from ***(5)***, follows: 50 mg (0.42 mmol/g, 1 eq) H-(BAL-Ile-MBHA-PS)-Leu-OOctyl (prepared as in **Subheading 3.1.1.**, but using 67 mg [10 eq] HCl•H-Leu-OOct instead of HCl•H-Leu-O*t*Bu) was washed with CH_2Cl_2 (2 × 0.5 min, 3 mL each). 81 mg (10 eq) Ddz-Phe-OH was dissolved in 0.5 mL CH_2Cl_2–DMF (9:1), 72 μL (20 eq) DIEA was added, the solution was added to the resin, and coupling was initiated by addition of 109 mg (10 eq) PyAOP as a solid. The coupling was allowed to proceed for 2 h. The resin was washed with CH_2Cl_2 (5 × 0.5 min, 3 mL each) and DMF (5 × 0.5 min, 3 mL each), and the coupling procedure was repeated for 2 h. The Ddz group was removed by treatment with TFA–H_2O–CH_2Cl_2 (3:1:966 × 1 min, 3 mL each) (with negligible loss of chains due to premature cleavage of the BAL anchor), followed by washing with CH_2Cl_2 (5 × 0.5 min, 3 mL each). Next, 63 mg (10 eq) Fmoc-Gly-OH and 109 mg (10 eq) PyAOP were dissolved separately in 0.4 mL DMF, combined, and added to the resin, and *in situ* neutralization/coupling initiated by the addition of 72 μL (20 eq) DIEA was carried out for 2 h. Finally, the resin was washed with CH_2Cl_2 (5 × 0.5 min, 3 mL each) and DMF (5 × 0.5 min, 3 mL each).

3.1.4. Stepwise Peptide Elongation Using Fmoc Protocols

With the C-terminal residue (*t*Bu ester or most other modified end groups) introduced as part of the BAL anchor, and the N^{α}-Fmoc protected penultimate

residue incorporated by the optimized acylation conditions described in **Subheading 3.1.2.**, further stepwise chain elongation by addition of Fmoc-amino acids proceeds normally using Fmoc protocols. However, when the C-terminal residue introduced by reductive amination (*see* **Subheading 3.1.1.**) has *n*-alkyl or allyl ester protection, DKP-avoiding protocols (*see* **Subheading 3.1.3.**) are necessary to successfully incorporate the second and third residues. There are quite a few ways to successfully carry out cycles of Fmoc SPS; a detailed experimental procedure for a specific example, taken again from ***(5)***, follows: 21.5 µmol Fmoc-Phe-(BAL-Ile-MBHA-PS)-Leu-O*t*Bu (prepared as in **Subheading 3.1.2.**) was washed with CH_2Cl_2 (2 × 0.5 min, 3 mL each), and DMF (2 × 0.5 min, 3 mL each). The resin was deprotected by treatment with piperidine–DMF (1:4, 3 × 1 min, 2 × 5 min, 3 mL each), followed by washes with DMF (5 × 0.5 min, 3 mL each). Next, 32.0 mg (5 eq) Fmoc-Gly-OH and 16.5 mg (5 eq) HOBt were dissolved separately in 0.4 mL DMF, combined, added to the resin, and coupling initiated by addition of 17 µL (5 eq) DIPCDI. Acylation was carried out for 1 h followed by washing the resin with DMF (5 × 0.5 min, 3 mL each), and CH_2Cl_2 (5 × 0.5 min, 3 mL each). The cycle was repeated to incorporate additional Fmoc-amino acids.

3.1.5. Acidolytic Cleavage to Release C-Terminal Modified Peptide From Resin

The BAL approach, as implemented with the acid-labile tris(alkoxy)benzylamide system, involves a single TFA cleavage step during which side-chain deprotection, C-terminus deprotection (as appropriate, for example, when cleavage of an acetal protecting group provides a C-terminal aldehyde functionality), and release of the final unprotected peptide into solution all occur at the same time. Several TFA/scavenger cocktails have been found to be suitable for Fmoc chemistry, (e.g., TFA–H_2O [19:1], TFA–Et_3SiH–H_2O [92:5:3], TFA–phenol–H_2O–Et_3SiH [88:5:5:2] ***[12]***, and TFA–thioanisole–1,2-ethanedithiol–anisole [90:5:3:2] ***[6]***), with generally good cleavage yields. A detailed experimental procedures for a specific example of cleavage, taken from ***(5,9),*** follows: 50 mg of the completed peptide-resin was cleaved with TFA–phenol–H_2O–Et_3SiH (88:5:5:2) (1 mL total) at 25°C for 2 h. The filtrate from the cleavage reaction was collected, combined with 1-mL TFA washes of the cleaved peptide-resin, concentrated under a stream of nitrogen, precipitated, and washed with diethyl ether (2 × 5 mL), dissolved in acetic acid (HOAc)–H_2O (1:4), and lyophilized.

3.2. BAL Strategy Modification I

The steps involved in this synthetic sequence (**Fig. 2**) are described in **Subheadings 3.2.1.–3.2.5.** These include: (1) reductive amination involving BAL-resin and the penultimate residue of the peptide target, with its carboxyl group

orthogonally protected; (2) peptide chain assembly in the $C \rightarrow N$ direction, following procedures described in **Subheadings 3.1.3.–3.1.4.**; (3) selective orthogonal removal of the carboxyl protecting group; (4) solid-phase activation of the pendant carboxylic acid and coupling with a C-terminal modified amino acid derivative; and (5) acidolytic cleavage to release the peptide into solution.

3.2.1. Reductive Amination to Anchor an Amino Acid Allyl Ester to BAL

This is done as previously described in **Subheading 3.1.1.**, but using, for example, 10 eq HCl•H-Ala-OAl instead of HCl•H-Leu-O*t*Bu.

3.2.2. Peptide Assembly in C → N Direction

Preparation of the C-terminal allyl ester protected peptide-resin is carried out by careful introduction of the second and third residues using protocols to minimize DKP formation followed by standard stepwise chain elongation as outlined in **Subheadings 3.1.3.** and **3.1.4.**, respectively.

3.2.3. Selective Orthogonal Removal of Allyl Ester Protecting Group

Anchoring an amino acid with an orthogonally removable C-terminal carboxyl protecting group (e.g., allyl, removed by treatment with palladium (0) in the presence of appropriate nucleophiles) to BAL-resin allows selective liberation of this carboxyl group for further manipulation while the peptide is bound to the solid support. The appropriate C-terminal modified amino acid derivative is coupled to the peptide in the $N \rightarrow C$ direction. A detailed experimental procedure for a specific example, taken from ***(9)***, follows: 50 mg (0.14 mmol/g) Boc-Val-Tyr(*t*Bu)-Phe(BAL-Ile-PEG-PS)Ala-OAl resin was washed with CH_2Cl_2 (5 × 0.5 min, 3 mL each) and DMF (5 × 0.5 min, 3 mL each), and the C-terminal allyl ester was cleaved by treatment with 40 mg (5 eq) $Pd(PPh_3)_4$ in $CHCl_3$–HOAc–NMM (37:2:1, 1.0 mL total) under argon at 25°C for 3 h. The peptide-resin was then washed with THF (3 × 2 min, 3 mL each), DMF (3 × 2 min, 3 mL each), CH_2Cl_2 (3 × 2 min, 3 mL each), DIEA–CH_2Cl_2 (1:19, 3 × 2 min, 3 mL each), CH_2Cl_2 (3 × 2 min, 3 mL each), sodium *N,N*-diethyldithiocarbamate (0.03 *M* in DMF, 3 × 15 min, 3 mL each), DMF (5 × 2 min, 3 mL each), CH_2Cl_2 (3 × 2 min, 3 mL each), and DMF (3 × 1 min, 3 mL each).

3.2.4. Solid-Phase Coupling With C-Terminal Species

The next synthetic step is activation of the pendant carboxyl group and subsequent condensation with the desired C-terminal modified amino acid derivative. A readily epimerizable oxazolonium ion may form during the activation step by attack of the oxygen from the BAL-amide function to the activated carboxyl; despite this risk, protocols have been developed to minimize racemization of this C-terminal residue ***(13)***. Features of these protocols include care-

ful optimization of activation reagents, solvents, bases, and temperatures, as well as absence of preactivation. Several unprotected *p*-nitroanilide and *S*-benzyl, *S*-phenyl, and *S*-alkyl thioester peptides have been prepared in excellent yields and purities using "BAL strategy modification I." A detailed experimental procedure for a specific example, taken again from ***(9)***, follows: A solution of 17 mg (10 eq) HCl•H-Ala-*p*NA in 0.28 mL CH_2Cl_2 and 24 µL (20 eq) DIEA was added to 50 mg (0.14 mmol/g) Boc-Val-Tyr(*t*Bu)-Phe(BAL-Ile-PEG-PS)Ala-OH resin; next, the coupling was initiated by addition of 27 mg (10 eq) HATU. After 30 min at 25°C, the peptide-resin was washed with DMF (6 × 0.5 min, 3 mL each) and CH_2Cl_2 (4 × 0.5 min, 3 mL each).

3.2.5. Acidolytic Cleavage

This proceeds as described previously in **Subheading 3.1.5.**

3.3. BAL Strategy Modification II

Synthetic steps are described in **Subheadings 3.3.1.–3.3.6.** (**Fig. 3**). These include (1) start of SPS by anchoring to BAL the penultimate residue of the peptide target, with its carboxyl group protected as an allyl or *tert*-butyl ester; (2) selective removal of the carboxyl protecting group, without premature cleavage of the BAL anchor; (3) solid-phase activation of the carboxyl group using a phosphonium salt and coupling with the desired C-terminal modified amino acid derivative; (4) acylation of the BAL-anchored amine; (5) Fmoc peptide chain elongation; and (6) TFA cleavage.

3.3.1. Reductive Amination Procedure

This is done as described in **Subheading 3.1.1.**, but using the eventual penultimate residue of the peptide target with its carboxyl group protected as an allyl or *tert*-butyl ester.

3.3.2. Selective Removal of Allyl or tert-Butyl Ester Protecting Groups

The next reaction for stepwise synthesis requires removal of the carboxyl-protecting group, without premature release of the penultimate residue from BAL-resin. Both allyl and *tert*-butyl ester protecting groups are compatible, with the latter providing simpler experimental procedures for removal and greater chemical diversity resulting from their commercial availability. The chemical bond between BAL-resin and the N^{α}-amino group of the penultimate residue is stable to the conditions required to remove the *t*Bu ester carboxyl-protecting group. (Once the amine is acylated, the peptide attachment to the resin becomes labile to high concentrations of TFA.) Allyl ester removal is described in **Subheading 3.2.3.** and a detailed experimental procedure for a specific example of *tert*-butyl ester removal, taken from ***(10)***, follows: 50 mg (0.83 mmol/g, 1 eq) H-

(BAL-Ile-MBHA-PS)-Ala-O*t*Bu was washed with CH_2Cl_2 (2 × 0.5 min, 3 mL each) and the C-terminal *tert*-butyl ester was cleaved by treatment with neat TFA (6 × 10 min, 3 mL each) at 25°C. The peptide-resin was then washed with CH_2Cl_2 (5 × 0.5 min, 3 mL each), DIEA–CH_2Cl_2 (1:19, 3 × 1 min, 3 mL each), CH_2Cl_2 (5 × 0.5 min, 3 mL each), and DMF (5 × 0.5 min, 3 mL each).

3.3.3. Solid-Phase Activation and Coupling With C-Terminal Species

The next reaction consists of activation of the carboxyl group using a phosphonium salt, and coupling with the desired C-terminal residue. In this approach, racemization during activation can occur only by a direct enolization mechanism, insofar as the oxazolone mechanism is precluded; indeed, racemization was essentially not observed. Additionally, aminium/uronium-coupling reagents should be avoided because of potential undesired guanidino formation between the coupling reagent and the BAL-anchored amine ***(14)***. A detailed experimental procedure for a specific example, taken again from ***(10)***, follows: A solution of 48 mg (10 eq) H-Sar-N$(CH_3)_2$ in 0.6 mL DMF and 141 µL (20 eq) DIEA was added to 50 mg (0.83 mmol/g, 1 eq) H-(BAL-Ile-MBHA-PS)-Ala-OH; next, the coupling was initiated by addition of 216 mg (10 eq) PyBOP as a solid. After 90 min at 25°C, the peptide-resin was washed with DMF (6 × 0.5 min, 3 mL each) and CH_2Cl_2 (4 × 0.5 min, 3 mL each).

3.3.4. Acylation of BAL-Anchored Amino Dipeptide

As previously described in **Subheading 3.1.2.**

3.3.5. Fmoc Peptide Chain Elongation

As previously described in **Subheading 3.1.4.**

3.3.6. Acidolytic Cleavage

Peptide *N*,*N*-dialkylamides containing prolyl or *N*-alkylamino acyl derivatives at the C-terminus can be obtained using BAL strategy modification II following acidolytic cleavage, as described previously in **Subheading 3.1.5.** However, the use of TFA–H_2O (19:1) as a cleavage cocktail is not recommended because partial hydrolysis (5–9%) of the amide bond attaching the penultimate residue bound to BAL and the *N*-alkyl C-terminal residue occurs. This side reaction can be prevented by using TFA–Et_3SiH (19:1) for cleavage.

4. Notes

1. BAL resin, formerly named *o,p*-PALdehyde-resin, is commercially available from Applied Biosystems, Framingham, MA.
2. When incorporating an optically active amino acid derivative, a separate imine-forming step should be avoided to reduce racemization.

3. Isoleucine (Ile) was used as an internal reference amino acid (IRAA) ***(15)***.
4. Manual SPS was carried out in polypropylene syringes (disposable reaction vessels) equipped with a porous polypropylene disk at the bottom. Syringes of variable volume can be used depending on the quantity of initial dried resin (e.g., 3.5-mL syringes for 100 mg of resin, 5-mL syringes for 200-mg of resin, and 20-mL syringes for 1 g of resin). Typically, resin was added to the syringe, and then the solvent used in the following reaction was added to create a slurry. The resin beads were washed with this solvent (3 mL of solvent per 1 mL of swollen resin). The mixture was stirred using a Teflon rod for a given time, and after finishing the treatment, the solvent was removed by filtration via aspiration. Prior to carrying out a reaction, the bottom part of the syringe was capped using a septum, and then solvents and reagents were added. After manual stirring using a Teflon® rod for 3 min, the mixture was allowed to react for a given time with occasional manual agitation.
5. For acylation of resin-bound secondary amines, the choice of solvent is critical. We found that CH_2Cl_2 or CH_2Cl_2–DMF (9:1) (the amount of DMF is needed for solubility reasons) gave the optimal results.

References

1. Kates, S. A. and Albericio, F. (eds.) (2000) *Solid-Phase Synthesis. A Practical Guide*. Marcel Dekker, New York.
2. Goodman, M., Felix, A., Moroder, L., and Toniolo, C. (eds.) (2002) *Houben-Weyl. Methods of Organic Chemistry. Synthesis of Peptides and Peptidomimetics,* vol. E22a. Georg Thieme Verlag, Stuggart-New York.
3. Songster, M. F. and Barany, G. (1997) Handles for solid-phase peptide synthesis, in *Methods in Enzymology, Solid-Phase Peptide Synthesis,* vol. 289 (Fields, G. B., ed.), Academic Press, Orlando, FL, pp. 126–174.
4. Alsina, J., Jensen, K. J., Albericio, F., and Barany, G. (1999) Solid-phase synthesis with tris(alkoxy)benzyl backbone amide linkage (BAL) *Chem. Eur. J.* **5,** 2787–2795.
5. Jensen, K. J., Alsina, J., Songster, M. F., Vágner, J., Albericio, F., and Barany, G. (1998) Backbone amide linker (BAL) strategy for solid-phase synthesis of *C*-terminal modified and cyclic peptides. *J. Am. Chem. Soc.* **120,** 5441–5452.
6. Albericio, F., Kneib-Cordonier, N., Biancalana, S., et al. (1990) Preparation and application of the 5-(4-(9-fluorenylmethyloxycarbonyl)aminomethyl-3,5-dimethoxyphenoxy)valeric acid (PAL) handle for the solid-phase synthesis of *C*-terminal peptide amides under mild conditions. *J. Org. Chem.* **55,** 3730–3743.
7. Guillaumie, F., Kappel, J. C., Kelly, N. M., Barany, G., and Jensen, K. J. (2000) Solid-phase synthesis of *C*-terminal peptide aldehydes from amino acetals anchored to a backbone amide linker (BAL) handle. *Tetrahedron Lett.* **41,** 6131–6135.
8. Brask, J., Albericio, F., and Jensen, K. J. (2003) Fmoc solid-phase synthesis of peptide thioesters by masking as trithioortho esters. *Org. Lett.* **5,** 2951–2953.
9. Alsina, J., Yokum, T. S., Albericio, F., and Barany, G. (1999) Backbone amide linker (BAL) strategy for N^{α}-9-fluorenylmethoxycarbonyl (Fmoc) solid-phase syn-

thesis of unprotected peptide *p*-nitroanilides and thioesters. *J. Org. Chem.* **64,** 8761–8769.

10. Alsina, J., Yokum, T. S., Albericio, F., and Barany, G. (2000) A modified backbone amide linker (BAL) solid-phase peptide synthesis strategy accommodating prolyl, *N*-alkylamino acyl, or histidyl derivatives at the *C*-terminus. *Tetrahedron Lett.* **41,** 7277–7280.
11. Alsina, J., Giralt, E., and Albericio, F. (1996) Use of *N*-tritylamino acids and PyAOP for the suppression of diketopiperazine formation in Fmoc/^{t}Bu solid-phase peptide synthesis using alkoxybenzyl ester anchoring linkages. *Tetrahedron Lett.* **37,** 4195–4198.
12. Solé, N. A. and Barany, G. (1992) Optimization of solid-phase synthesis of [Ala^{8}]-dynorphin-A. *J. Org. Chem.* **57,** 5399–5403.
13. Lloyd-Williams, P., Albericio, F., and Giralt, E. (1997) *Chemical Approaches to the Synthesis of Peptides and Proteins*. CRC, Boca Raton, FL.
14. Albericio, F., Bofill, J. M., El-Faham, A., and Kates, S. A. (1998) Use of onium salt-based coupling reagents in peptide synthesis. *J. Org. Chem.* **63,** 9678–9683.
15. Albericio, F. and Barany, G. (1985) Improved approach for anchoring N^{α}-9-fluorenylmethyloxycarbonylamino acids as *p*-alkoxybenzyl esters in solid-phase peptide synthesis. *Int. J. Peptide Protein Res*. **26,** 92–97.

13

Synthesis of Peptide Bioconjugates

Ferenc Hudecz

Summary

Bioconjugates play an important role in several fields of biomolecular and biomedicinal sciences. Protein/polypeptide-based conjugates with covalently attached epitope peptides are considered as potential synthetic vaccine candidates and/or target antigens in affinity-based bioassays. This chapter describes the synthesis of two- and three-component bioconjugates using water-soluble branched chain polymeric polypeptides with multiple amino and/or carboxyl groups as macromolecular partners and oligopeptides as epitopes with small molecular mass. The synthetic procedures outline three major strategies for the incorporation of multiple copies of uniformly oriented peptide epitopes. In the first example, chloroacetylated polypeptide is conjugated with SH-peptide to form a thioether linkage. Second, two independent oligopeptides are introduced into a macromolecule by amide and disulfide bonds, respectively. In the third example, a new procedure is reported for the formation of disulfide bridges by the use of Npys-modified polypeptide and SH-peptide.

Key Words: Peptide–macromolecule conjugate; amide bond; thioether bond; disulfide bridge; antibody epitope; T-cell epitope peptide.

1. Introduction

Bioconjugates play an important role in several fields of biomolecular and biomedicinal sciences. Protein/polypeptide-based conjugates with covalently attached epitope peptides are considered as potential synthetic vaccine candidates and/or target antigens in affinity-based bioassays (e.g., ELISA, BIACORE) ***(1,2)***. Drugs such as daunomycin, methotrexate, and a GnRH analog coupled either to monoclonal antibodies or to synthetic linear or branched polypeptides

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

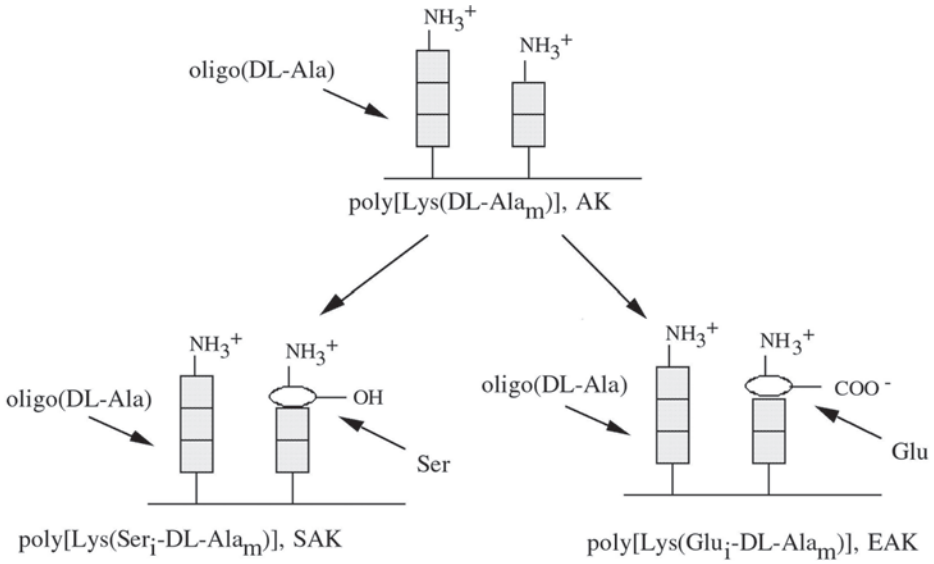


Fig. 1. Schematic representation of branched-chain polymeric polypeptides.

exhibit altered biodistribution and decreased nonspecific toxicity (e.g., cardiotoxicity). Consequently, the therapeutic effects of such drugs can be markedly improved by conjugation ***(3–5)***. Bioconjugates could be classified either by the number and size of components involved or by the type of covalent linkage applied between the partner entities. This chapter describes the synthesis of two- and three-component bioconjugates combining water-soluble branched-chain polymeric polypeptides as the macromolecular partners of small molecular mass bioactive oligopeptides. Collectively, the bioconjugates presented possess not only the most frequently utilized amide and/or disulfide linkage, but also the thioether bond. The synthetic procedures outlined could also be adapted to the preparation of conjugates containing macromolecular partners with similar functional groups (e.g., proteins and synthetic polymers).

Branched polypeptides developed in our laboratory with the general formula poly[Lys(X_i-DL-Ala$_m$)] (XAK), where $i < 1$, m approx 3, and X represents the side-chain terminal residue ***(6–8)*** will be used (**Fig. 1**) to illustrate the methods utilized for different conjugation procedures. Depending on the nature of the amino acid X, polypeptides exhibit polycationic (e.g., poly[Lys(DL-Ala$_m$)], poly[Lys(Ser$_i$-DL-Ala$_m$)], amphoteric (e.g., poly[Lys(Glu$_i$-DL-Ala$_m$), or polyanionic (e.g., poly[Lys(Ac-Glu$_i$-DL-Ala$_m$)]) character under physiological conditions (pH 7.3 in 0.15 *M* NaCl).

These macromolecules were applied for the synthesis of B-cell epitope peptide conjugates to be used as target antigens for the specific and sensitive detection of mucin 1 glycoprotein-specific antibodies ***(9)***, of T-cell epitope conjugates with peptides derived from 16/38-kDa proteins from *M. tuberculosis* ***(10–13)*** and of peptides containing HSV-neutralizing antibody epitope(s) ***(14,15)***. In these conjugates the oligopeptide component(s) was/were coupled in multiple copies in a uniformly oriented manner. By these constructs we have shown that (1) peptides attached to carriers preserved their biological function (e.g., antibody binding, T-cell recognition, specific immunogenicity, interaction with phospholipid mono- or bilayers), and (2) the physicochemical properties of the polypeptide component have a marked influence on the attached epitope-related activities ***(1,13,16)***.

2. Materials

1. Amino acid derivatives, reagents for coupling and analysis, organic solvents, buffers, and HPLC eluents used for the preparation of oligo- and polypeptides are identical with those described in Chapter 4.
2. Branched-chain polypeptides poly[Lys-(X_i-DL-Ala$_m$)], where X = Glu, Ser, or Boc-Cys(Npys)Opfp, were produced in our laboratory as described.
3. Cross-linking compound]3-(2-pyridyldithio) propionic acid *N*-hydroxy-succinimide ester (SPDP) (Sigma, Poole, UK).
4. Benzotriazol-1-yl-oxytris(dimethylamino)phosphonium hexafluorophosphate (BOP) (Fluka, Buchs, Switzerland).
5. Chloroacetic acid pentachlorophenyl ester (ClAc-Opcp) (Fluka).
6. DL-dithiothreitol (DTT) (Fluka).

3. Methods

Abbreviations used in this chapter follow the rules of the IUPAC-IUB Commission of Biochemical Nomenclature ***(17)*** in accordance with the recommended nomenclature of graft polymers ***(18)***.

3.1. Synthesis of Branched-Chain Polymer Polypeptides

The synthesis of branched polypeptides used for the preparation of all conjugates described in this account was published in detail earlier ***(6–8)***. Briefly, poly[Lys] was prepared by the polymerization of N^{α}-carboxy-N^{ε}-benzyloxycarbonyl-lysine anhydride. After cleavage of the protecting groups, poly[Lys (DL-Ala$_m$)] (AK) was produced by grafting of short oligomeric DL-Ala side chains onto the ε-amino groups of poly[Lys]. Benzyloxycarbonyl-protected amino acid derivatives (Z-Ser-OPcp and Z-Glu(OBzl)-OPcp) were coupled to the end of the side chains of AK by the HOBt-catalyzed active ester method. Blocking groups were removed completely with HBr in glacial acetic acid, as

confirmed by UV spectroscopy at 254 nm, resulting in poly[Lys(X_i-DL-Ala$_m$)] (XAK) polymers (where i ≤ 1 and m ≈ 3, *X* = Ser or Glu). The amino acid composition of polypeptides as well as peptide conjugates were determined by amino acid analysis using a Beckman (Fullerton, CA) model 6300 amino acid analyzer. Prior to the analysis, samples were hydrolyzed in 6 *M* HCl in sealed and evacuated tubes at 110°C for 24 h. The size and molecular weight of polymers were calculated from the amino acid composition and the sedimentation analysis of poly[Lys] ***(19)***.

3.2. Synthesis and Characterization of Epitope Peptides

Peptides used in these studies for conjugate synthesis were prepared by Boc/Bzl strategy on MBHA resin or by Fmoc-based chemistry on Rink-amide resin as described in Chapter 4. For incorporation of SH group at the N-terminal (C^{91}SEFAYGSFVRTVSLPVGADE^{110}) or C-terminal position (H-^{9}LKNleADPNRFRGKDL^{22}C-NH_2, [Nle^{11}]-9-22-Cys) or in a central position (H-LKNleADPNRFRGKDL-Acp-CSALLEDPVG-NH_2), a Cys derivative, Boc-Cys(Meb)-OH, was added to the native sequence. After the removal of the peptides from the resin the crude products were purified by RP-HPLC and the purified samples were characterized by analytical HPLC, amino acid analysis and mass spectrometry as detailed in Chapter 4.

3.3. Two-Component Conjugate Containing Multiple Copies of Uniformly Oriented Antibody Peptide Epitope

Synthetic oligopeptides comprising linear or continuous topographic B-cell epitope sequences might be considered as specific and small-size antigens. It has been demonstrated that the avidity and specificity of antibody binding could be altered by conjugation to macromolecules or by modification in the flanking regions. However, no systematic studies have been reported to describe the effect of different carrier macromolecules in epitope conjugates. To this end the influence of carrier structure and topology on the antibody recognition of covalently attached herpes simplex virus type 1 glycoprotein D (HSV-1 gD)-related epitopes have been studied by comparing the monoclonal antibody binding properties of a new set of conjugates ***(20)*** with branched-chain polypeptide, poly[Lys(Ser$_i$-DL-Ala$_m$)] (SAK), tetratuftsin analog (H-[Thr-Lys-Pro-Lys-Gly]$_4$-NH_2) sequential oligopeptide carrier (SOC$_n$), multiple antigenic peptide (MAP), and keyhole limpet hemocyanine (KLH). In these novel constructs peptide ^{9}LKNleADPNRFRGKDL^{22} ([Nle^{11}]-9-22) representing an immunodominant B-cell epitope ***(21,22)*** was conjugated to polypeptides through thioether bonds (**Fig. 2**).

First, we prepared a derivative of polypeptide SAK possessing chloroacetyl groups in multiple copies at the α-amino groups of the branches. This compound

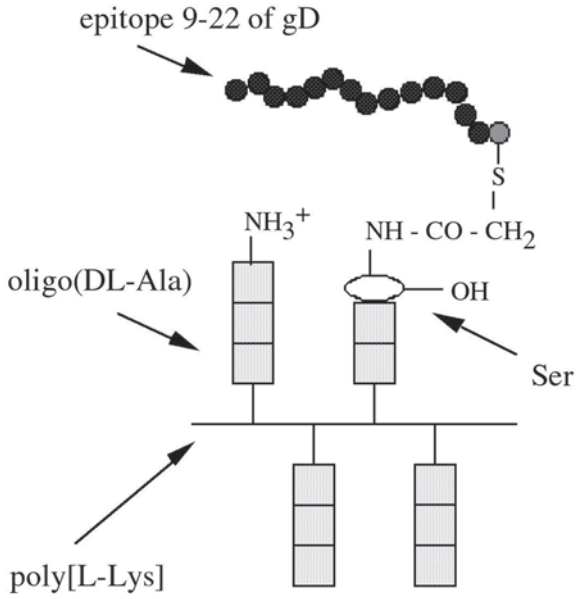


Fig. 2. Schematic structure of branched polypeptide (SAK) conjugate with multiple copies of peptide epitope corresponding to the 9-22 region of glycoprotein D of HSV-1.

was conjugated with epitope peptides possessing Cys at C-terminal position as outlined in **Fig. 3**. Binding data suggest that the chemical nature of the carrier and the degree of substitution have marked influence on the avidity of antibody binding.

3.3.1. Synthesis of Chloroacetylated-Branched Polypeptide poly[Lys(ClAc$_j$-Ser$_i$-DL-Ala$_m$)] (SAK(ClAc)$_j$)

80 mg (0.145 mmol) poly[Lys(Ser$_{0.9}$-DL-Ala$_{3.5}$)] HBr salt (SAK, DP_n = 60, $MW_{monomer}$ = 550) was dissolved in 1 mL deionized water and the solution was diluted with 4 mL DMF. Chloroacetyl groups were introduced at the N^{α}-amino group of Ser residues by chloroacetic acid pentachlorophenyl ester. Six different SAKs—ClAc-OPcp ratio (1:1, 1:0.8, 1:0.6, 1:0.5, 1:0.4 and 1:0.3 mol/mol; 50 mg, 40 mg, 30 mg, 25 mg, 20 mg, 15 mg ClAc-OPcp respectively) were used and the active ester was added to the polymer containing solution dissolved in 5 mL DMF. The reaction mixtures were stirred overnight at room temperature. The solution was filled in Visking tubes (cutoff 8000–12,000) and dialyzed for 2 d against 0.1% acetic acid and freeze-dried. According to Cl analyses, 46.5%, 45.9%, 48.5%, 41.3%, 30.1%, and 21.7% of the side chains were blocked by the ClAc group, respectively.

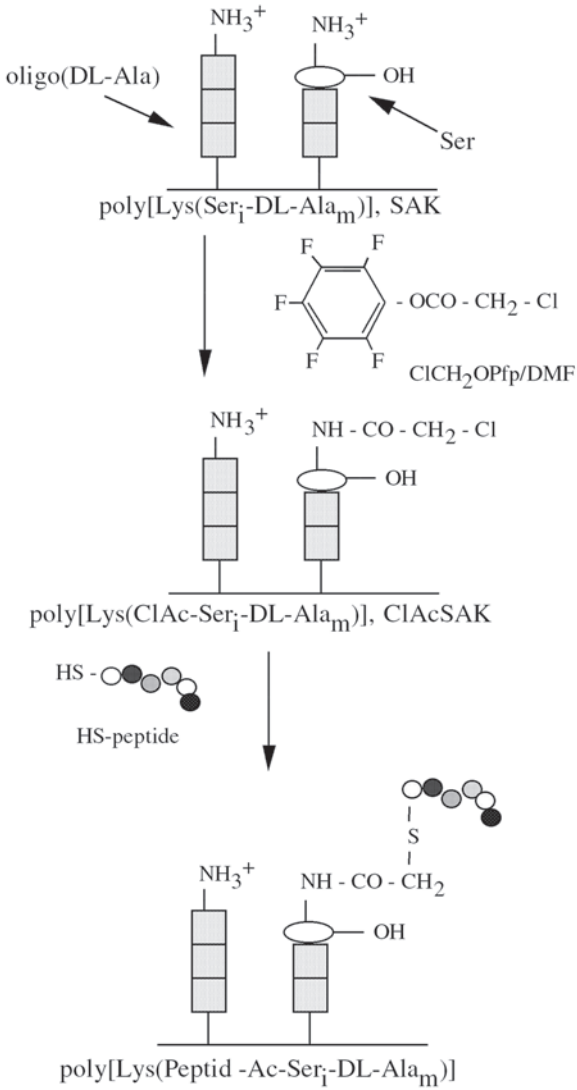


Fig. 3. Outline of the synthesis of ([Nle[11]]-9-22-Cys)-conjugate with chloroacetylated SAK by introduction thioether bond.

3.3.2. Conjugation of Peptide H-[9]LKNleADPNRFRGKDL[22]C-NH_2 ([Nle[11]]-9-22-Cys) With Chloroacetylated Branched Polypeptide poly[Lys($ClAc_j$-Ser_i-DL-Ala_m)] (SAK$(ClAc)_j$)

For the preparation of multivalent epitope conjugates, branched-chain polymeric polypeptide (SAK) was reacted with chloroacetic acid pentachlorophe-

nyl ester in DMF-water (9:1, v/v) solution. Four differently substituted carriers were applied. 15 mg of SAK$(ClAc)_{48.5}$, SAK$(ClAc)_{41.3}$, SAK$(ClAc)_{30.1}$, or SAK $(ClAc)_{21.7}$ (acetate salt form) were dissolved in 100 mL 0.1 *M* Tris-HCl buffer (pH 8.2). 1.2 eq (calculated for the chloroacetyl content of the polymers) of epitope peptide (23.4 mg, 20 mg, 14.8 mg, and 10.6 mg respectively) were added in two portions to the solution. The conjugation reaction continued for 24 h and was terminated by addition of an excess of Cys to block the unreacted chloroacetyl groups. Crude products were dialyzed against water to remove uncoupled peptide and Cys in Visking tubes (cutoff 8000–12,000) for 2 d. The average degree of substitution was calculated from the amino acid analyses. Depending on the input molar ratio, 44%, 22%, 9%, and 7% of the side chains was substituted by epitope peptide, respectively.

3.4. Three-Component Conjugate Containing Multiple Copies of Two Different, But Uniformly Oriented, T-Cell Peptide Epitopes

For the preparation of three-component conjugates containing multiple copies of two independently introduced and uniformly oriented T-cell peptide epitopes, an amphoteric branched chain polypeptide, poly[Lys-(Glu_i-DL-Ala_m)] (EAK), was used. This polypeptide, with free α-amino and γ-carboxyl groups at the end of the side chains, was conjugated with peptides representing two immunodominant regions of the 16-kDa and 38-kDa proteins of *Mycobacterium tuberculosis*, respectively ***(10–12)***. For conjugation, peptide ^{65}FNLWGPA FHERYPNVTITA^{83} (peptide 1) from the 38-kDa protein possessing free α-amino group at the N-terminal was used without modification, while the second epitope peptide originating from the 16-kDa protein was elongated with an *N*-terminal Cys [C^{91}SEFAYGSFVRTVSLPVGADE^{110}] (peptide 2). Consequently, an amide between the γ-COOH of EAK and α-NH_2 of peptide 1 and a disulfide bridge between the α-NH_2 of EAK and SH of peptide 2 were introduced (**Fig. 4**).

This synthetic strategy is outlined in **Fig. 5**. In the first step, the α-aminogroup of Glu in the side chain of EAK was modified by the heterobifunctional reagent SPDP to introduce protected SH groups into the polymer structure of [(SSP)EAK] ***(23)***. The extent of 2-pyridyl-disulfide group incorporation was determined spectrophotometrically from the amount of pyridine-2-thione released by reduction with DTT ***(10)***. The unmodified, free α-amino groups of Glu in the branches were blocked by acetylation using a mixture of acetic anhydride and imidazole. In the next step, peptide 1 was coupled to (Ac,SSP)EAK. This was achieved by the BOP reagent-based activation method in which the γ-carboxyl group of glutamic acid of (Ac,SSP)EAK was linked to the α-amino group of the phenylalanine residue in peptide 1 to provide isopeptide (γ,α type amide) bonding. After isolation of the product the presence of the 2-pyridyl-

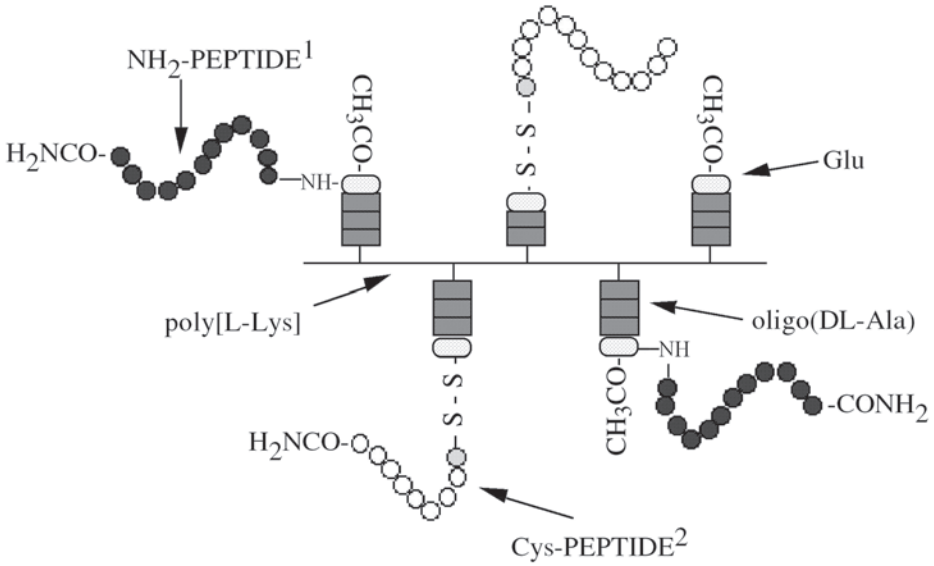


Fig. 4. Schematic structure of branched polypeptide (EAK) conjugate containing two different epitope peptides from *M. tuberculosis* proteins. ^{65}FNLWGPAFHERYPN VTITA^{83} (peptide 1) and C^{91}SEFAYGSFVRTVSLPV-GADE^{110} (peptide 2) are attached by amide and disulfide bonds, respectively.

disulfide groups was verified. In the last step, a disulfide bridge was introduced by the substitution of the protecting group of SH of the polymer by peptide 2 [C^{91}SEFAYGSFVRTVSLPVGADE^{110}]. The completion of the reaction between (Ac,SSP)E(peptide 1)AK and peptide 2 was spectrophotometrically determined.

In vitro T-cell immunogenicity data obtained with this conjugate using T cell hybridomas, lymph node cells from immunized mice, and human PBMC cultures from PPD-positive individuals indicated that both epitopes were efficiently recognized ***(10)***.

3.4.1. Synthesis of poly[Lys-(Ac$_y$,SSP$_x$-Glu$_i$-DL-Ala$_m$)], (Ac,SSP)EAK

100 mg (215 µmol) EAK was dissolved in 7.5 mL 0.1 *M* PBS (pH 8.0, adjusted with 1 *M* NaOH). 20 mg (64.0 µmol) SPDP (Sigma, Poole, UK) was dissolved in 2 mL abs. methanol, and added dropwise to the solution of the polypeptide. The reaction mixture was stirred for 30 min at room temperature and dialyzed for 24 h against distilled water (*see* **Note 1**).

110 mg (235 µmol) (SSP)EAK was dissolved in 1 mL distilled water and diluted to 5 mL with DMF at 4°C. In the meantime, 312 µL (3.3 mmol) acetic anhydride and 224 mg (3.3 mmol) imidazole were mixed in 1 mL DMF at 4°C

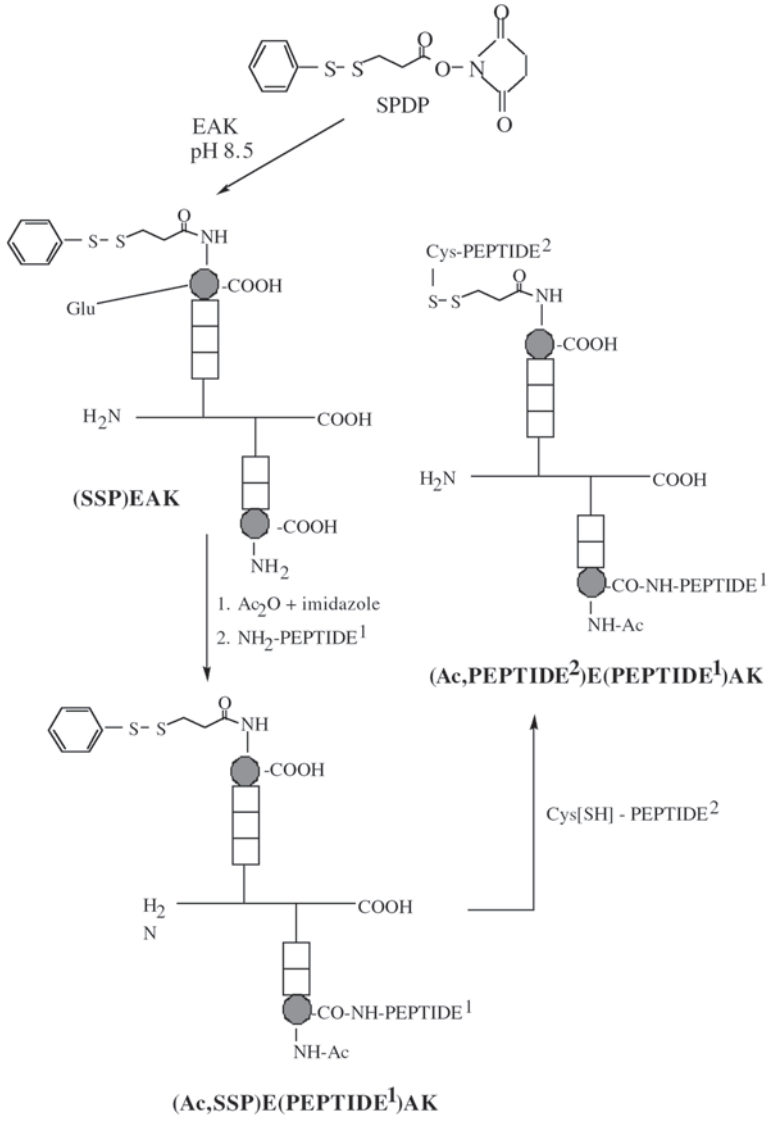


Fig. 5. Outline of the synthesis of conjugates containing two different epitope peptides, ^{65}FNLWGPAFHERYPNVTITA^{83} (peptide 1) and C^{91}SEFAYGSFVRTVSLPV GADE^{110} (peptide 2), and branched polypeptide EAK.

and stirred for 10 min. The two solutions were then mixed, and acetylation was continued for 2 h at room temperature. After dialysis against distilled water, the product was isolated by freeze drying (*see* **Note 2**).

3.4.2. Poly[Lys-(Ac_y,SSP_x)-Glu_i($Peptide_1$)-DL-Ala_m)], (Ac,SSP)E($Peptide_1$)AK

10 mg (19 μmol) (Ac,SSP)EAK (calculated on the basis of 19.3% 2-pyridyldisulfide group content, and 80.7% acetyl group content) was dissolved in 1 mL water and diluted five times with DMF. 8.9 mg (20 μmol) BOP reagent, 3.3 mg (25 μmol) HOBt, and 7 μL (40 μmol) DIEA were dissolved in 1 mL DMF. The activating mixture was added to the polymer solution and stirred for 20 min. 10 mg (4.5 μmol) peptide ^{65}FNLWGPAFHERYPNVTITA^{83} (peptide 1), representing 0.25 mol peptide/polymer side chain, was dissolved in 4 mL DMF at a concentration of 2.5 mg/mL DMF, and mixed with 4.5 μmol DIEA. The peptide solution was added to the preactivated polymer solution and allowed to couple for 24 h at room temperature under stirring. The solution was then dialyzed against distilled water for 48 h and the product was isolated by freeze drying.

3.4.3. Poly[Lys-(Ac_y,$Peptide_2$)-Glu_i($Peptide_1$)-DL-Ala_m)], (Ac,$Peptide_2$)E($Peptide_1$)AK

10 mg (Ac,SSP)E(Peptide1)AK conjugate (containing 3 μmol protected thiol groups) was dissolved in 1 mL distilled water. 10 mg (4.5 μmol) peptide C-^{91}SEFAYGSFVRTVSLPVGADE^{110} (peptide 2) dissolved in 1 mL PBS (pH 8.0, adjusted with 1 *M* NaOH) was added to the conjugate solution. The reaction mixture was stirred for 30 min at room temperature, followed by dialysis against distilled water for 48 h, and was freeze dried.

3.4.4. Analysis of Conjugates by RP-HPLC

The compositions of the conjugates were analysed by RP-HPLC with a Delta-Pak RP C_{18} column (7.8 mm × 30 cm) packed with spherical 15-μm silica of 300 Å pore size (Nixon Waters Ltd., Tokyo, Japan). We used an eluent system of 0.1% TFA in water (eluent A) and 0.1% TFA in acetonitrile/water = 80/20 v/v (eluent B) using high-purity (Analar) solvents and distilled, deionized water. The elution gradient was 10–100% eluent B in 35 min. The injection volume was 100 μL containing 10 μg of conjugate, polymer, or peptide. Compounds were dissolved in eluent immediately before application and filtered through 0.45-μm Spartan 13 (Schleicher and Schuell, Dassel, Germany) filters. Free peptide or polymer polypeptide samples were run as standards and their retention times were determined. UV absorbance was monitored at a wavelength of 220 nm. All analyses were carried out at ambient temperature with a flow rate of 1.0 mL/min. As a control, in some cases, the HPLC profile of the peptide and conjugate mixture was also recorded.

3.5. Three-Component Conjugate Containing Multiple Copies of Two Uniformly Oriented Antibody Peptide Epitopes

We have described a novel approach in which the carrier polypeptide is modified by 3-nitro-2-pyridinesulfenyl (Npys)-protected Cys ***(24)***; this derivative has been used for the conjugation of Cys-containing epitope peptides to poly [Lys]-based branched polypeptides ***(25)***. Considering the stability of the Npys group in the presence of pentafluorophenol, the Boc-Cys(Npys)-OPfp derivative was selected for the introduction of Npys groups to the N-terminal of the branches of polypeptide backbones. This new class of Cys(Npys)-derivatized branched polypeptides is stable for a couple of months and suitable for the effective preparation of epitope-peptide conjugates possessing increased water solubility. The incorporation of epitope peptides depends on the number of Npys groups in polymers as well as on the presence/absence of Boc-protecting group on the Cys residue. A conjugate was prepared in which two epitopes (^{9}LKNleA DPNRFRGKDL^{22} and ^{276}SALLEDPVG^{284}) ***(21,26)*** of HSV gD type 1 joined by a dipeptide Acp-Cys spacer (where Acp is e-amino caproic acid) were attached through the SH function of a single peptide to the Cys(Npys) modified branched polypeptide (**Fig. 6**).

3.5.1. Synthesis of Branched Polypeptides With Cys(Npys)Residues

3.5.1.1. Synthesis of Boc-Cys(Npys)-OPfp

500 mg (1.33 mmol) Boc-Cys(Npys)-OH (Mw: 375) and 245 mg (1.33 mmol) pentafluorophenol were dissolved in 5 mL DCM. Then 274 mg (1.33 mmol) DCC dissolved in 2 mL DCM was added to the solution and the reaction mixture was stirred in an ice bath for 15 min, then for 2 h at RT. The reaction was followed by t.l.c. DCU was filtered out after cooling the solution to 0°C and the solvent was evaporated. The crude product was purified by crystallization from methanol. Yield: 594 mg (83%) (Mw: 541). Melting point: 96–99°C. R_f: 0.92 (ethyl acetate), R_f: 0.65 (chloroform-methanol= 3:1, v/v). Optical rotation: $[\alpha]_D$ = −62.2 (c = 1, DMF).

3.5.1.2. The Synthesis of Poly[Lys(Cys(Npys)$_f$-DL-Ala$_M$)] (CAK)

20 mg (57 µmol) poly[Lys(DL-Ala$_{3.11}$)] (AK, DP_n = 66, $MW_{monomer}$ = 349) dissolved in 0.5 mL deionized water was diluted with 3 mL DMF. An equivalent amount of Boc-Cys(Npys)-OPfp (31 mg) dissolved in 2 mL DMF was added to the solution (*see* **Notes 3** and **4**). The reaction mixture was stirred overnight at RT. The solvent was removed *in vacuo* (30°C) and the remaining product was washed with ether (containing 10% DCM) repeatedly to remove unreacted ester and pentafluorophenol. Yield: 35 mg (*see* **Note 5**). The Boc-group was

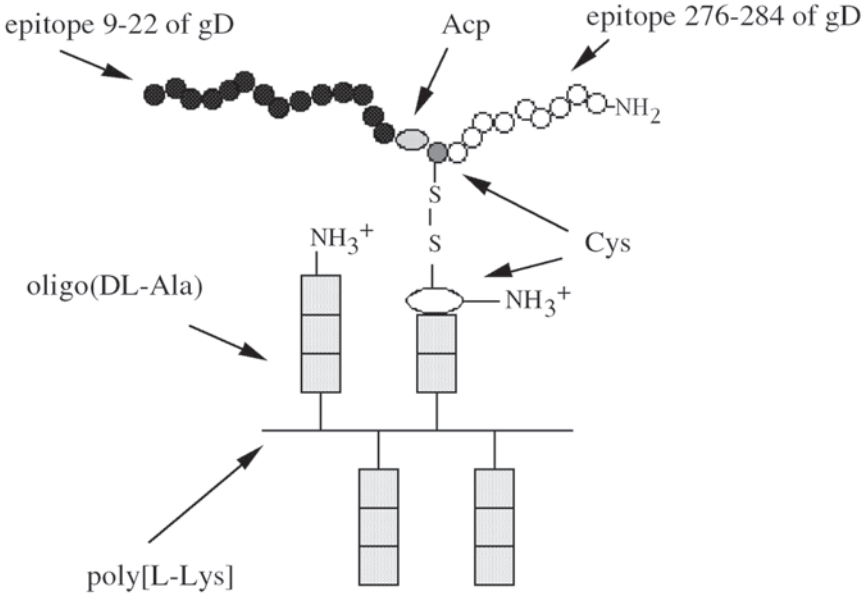


Fig. 6. Schematic structure of branched polypeptide (AK) conjugate with multiple copies of a single peptide containing epitopes corresponding to the 9–22 and to 276–284 regions of glycoprotein D of HSV-1. Two epitopes was joined first by an Acp-Cys dipeptide spacer and the SH function of the peptide was used for conjugation with Cys(Npys) with modified polypeptide.

removed by 5 mL TFA-water (95:5, v/v) at 0°C for 1 h in the presence of 5% *p*-cresol. The solution was concentrated *in vacuo*, then dissolved in water and dialyzed against 0.1% acetic acid solution for 2 d and freeze dried (yield: 27 mg). The average degree of substitution was calculated from the UV absorption of the Npys group at λ = 350 nm (ε = 3930/M/cm in DMF) (***27***).

3.5.2. Conjugation of Epitope Peptides to Branched Polypeptides With Cys(Npys) Residues

Boc-deprotected branched polypeptide with the Cys(Npys) moiety, poly [Lys($Cys_{0.27}$-DL-$Ala_{3.1}$)] trifluoroacetate salt (10 mg, 16.5 µmol), was dissolved in 10 mL 0.06 *M* phosphate buffer (pH 5.5) (*see* **Note 6**) 1.35 amount of H-LKNleADPNRFRGKDL-Acp-CSALLEDPVG-NH_2 (14.8 mg, 5.4 µmol), calculated from the Npys-content of polymer, was added to the reaction mixture. The final concentration was 1 mg/mL for the polymer component. The solution was stirred for 4 h at RT, transferred to a Visking tube (cutoff 8000–12,000), and dialyzed for 2 d against 0.1% acetic acid. No Npys content was detected by UV spectroscopy at λ = 350 nm in the conjugate after freeze drying (*see* **Note 7**).

The yield was 15 mg (86%). According to the amino acid analysis, 26% of side chains of branched polypeptide contained the bifunctional epitope peptide.

4. Notes

1. To determine the degree of 2-pyridyl-disulfide group incorporation, 400 µL of (SSP) EAK or (Ac,SSP)E(peptide1)AK solution was reacted with 200 µL DL-dithiothreitol (DTT) (2.5 mg/mL PBS), and after 5 min the absorbance of the released pyridine-2-thione was measured at λ = 343 nm, ε = 8080/M/cm ***(27)***.
2. The blocking of free amino groups by acetylation was verified by the ninhydrin assay ***(28)***.
3. Prior to the incorporation of protected Cys residue into polymers, the stability of the Npys group was studied in the presence of pentafluorophenol. This was caused by the formation of this compound as byproduct during the reaction between Boc-Cys(Npys)-OPfp and the α-amino groups of the N-terminal amino acids of branched polypeptides. In a model experiment, Boc-Cys(Npys)-OH was kept in DMF-water (9:1, v/v) mixture in the absence or presence of pentafluorophenol. Based on HPLC analysis we found that the Npys group is stable under these conditions.
4. The input Boc-Cys(Npys)-OPfp/polymer ratio should be optimized in each case. Depending on the macromolecular partner we found optimal incorporation in the range of 0.75–1.5:1 molar ratio.
5. Branched polypeptides containing Boc-Cys(Npys) can be dissolved in DMF (A) or in DMF-0.06 *M* phosphate buffer (pH 5.5) = 75:25 (v/v) mixture (B).
6. We observed a low level of SH-peptide substitution when Boc-protected branched polypeptides were used. This might be due to the steric hindrance caused by the presence of the large number of Boc-Cys(Npys) groups at the end of the side chains. Therefore, we selected the unprotected form of modified polymers containing 27% Cys(Npys). In these cases the average degree of substitution of branched polymers with the 25-mer peptide having a Cys residue in central position was 26%.
7. In case of the presence of UV band at λ = 350 nm, Npys groups were reacted with the 1.2 eq excess of Cys in 0.06 *M* phosphate buffer (pH 5.5) followed by dialysis and freeze-drying.

Acknowledgments

Experimental work summarized in this paper was supported by grants from WHO (T9/181/133), from the Hungarian Research Fund (OTKA T-014964, T-03838, T-043576), from the Ministry of Education (FKFP 0101/1997), from the Hungarian-Spanish Intergovernmental Programme (E-5/1998, E-3/2001), and from the "Peptide based synthetic antigens against infectious diseases" EU COST Chemistry Action (D13/0007/00).

References

1. Hudecz, F. (2001) Manipulation of epitope function by modification of peptide structure: a minireview. *Biologicals* **29,** 197–207.

2. Zeng, W., Ghosh, S., Macris, M., Pagnon, J., and Jackson, D. C. (2001) Assembly of synthetic peptide vaccines by chemoselective ligation of epitopes: influence of different chemical linkages and epitope orientations on biological activity. *Vaccine* **19,** 3843–3852.
3. Mezö, G., Mezö, I., Pimm, et al. (1996) Synthesis, conformation, biodistribution and hormon related in vitro antitumor effect of GnRH antagonist branched polypeptide conjugate. *Bioconjugate Chemistry* **7,** 642–650.
4. Hudecz, F., Kóczán, G., and Reményi, J. (2003) Peptide or protein based delivery and targeting, in *Molecular Pathomechanisms and New Trends in Drug Research* (Keri, G. and Toth, I., eds.), Taylor and Francis Group, London, pp. 553–578.
5. Hudecz, F., Reményi, J., Szabó, R., et al. (2003) Drug targeting by macromolecules without recognition unit? *J. Mol. Recognition* **16,** 288–298.
6. Hudecz, F. and Szekerke, M. (1980) Investigation of drug-protein interactions and the drug-carrier concept by the use of branched polypeptides as model systems. Synthesis and characterization of the model peptides. *Coll. Czech. Chem. Commun.* **45,** 933–940.
7. Mezö, G., Kajtár, J., Nagy, I., Szekerke, M., and Hudecz, F. (1997) Carrier design: Synthesis and conformational studies of poly[L-lysine] based branched polypeptides with hydroxyl groups. *Biopolymers* **42,** 719–730.
8. Hudecz, F., Pimm, M. V., Rajnavölgyi, É., et al. (1999) Carrier design: New generation of polycationic branched polypeptides containing OH groups with prolonged blood survival and diminished in vitro cytotoxicity. *Bioconjugate Chemistry* **10,** 781–790.
9. Hudecz, F. and Price, M. R. (1992) Monoclonal antibody binding to peptide epitopes conjugated to synthetic branched polypeptide carriers. Influence of the carrier upon antibody recognition. *J. Immunol. Methods* **147,** 201–210.
10. Wilkinson, K. A., Vordermeier, M. H., Wilkinson, R., Iványi, J., and Hudecz, F. (1998) Synthesis and in vitro T cell immunogenicity of conjugates with dual specificities: attachment of epitope peptides of 16 kDa and 38 kDa proteins from M. tuberculosis to branched polypeptide. *Bioconjugate Chemistry* **9,** 539–547.
11. Vordermeier, H. M., Harris, D. P., Roman, E., Lathigra, R., Moreno, C., and Ivanyi, J. (1991) Identification of T-cell stimulatory peptides from the 38 kDa protein of *M. tuberculosis*. *J. Immunol.* **147,** 1023–1029.
12. Friscia, G., Vordermeier, H. M., Pasvol, G., Harris, D. P., Moreno, C., and Iványi, J. (1995) Human T cell responses to peptide epitopes of the 16-kD antigen in tuberculosis. *Clin. Exp. Immunol.* **102,** 53–57.
13. Wilkinson, K. A., Hudecz, F., Vordermeier, H. M., Iványi, J., and Wilkinson, R. J. (1999) Enhancement of the T cell response to a mycobacterial peptide by conjugation to synthetic branched polypeptide. *Eur. J. Immunol.* **29,** 2788–2796.
14. Hilbert, Á., Hudecz, F., Mezõ, G., et al. (1994) The influence of branched polypeptide carriers on the immunogenicity of predicted epitopes of HSV-1 glycoprotein D. *Scand. J. Immunol.* **40,** 609–617.
15. Mezö, G., Dalmadi, B., Mucsi, I., Bösze, S., Rajnavölgyi, É., and Hudecz, F. (2002) Peptide based vaccine design: Synthesis and immunological characterisa-

tion of branched polypeptide conjugates comprising the 276-284 immunodominant epitope of HSV-1 glycoprotein D. *J. Peptide Science* **8,** 107–117.
16. Hudecz, F., Nagy, I. B., Kóczán, G., Alsina, M. A., and Reig, F. (2001) Carrier design: influence of charge on interaction of branched polymeric polypeptides with phospholipid model membranes, in *Biomedical Polymers and Polymer Therapeutics* (Chiellini, E., Sunamoto, J., Migliaresi, C., Ottenbrite, R. M., and Cohn, D. eds.), Kluwer Academic/Plenum Publishers, New York, pp. 103–120.
17. IUPAC-IUB Commission on Biochemical Nomenclature. (1972) *Biochem. J.* **127,** 753–756.
18. IUPAC-IUB Commission on Biochemical Nomenclature. (1984) *Eur. J. Biochem.* **138,** 9–37.
19. Hudecz, F., Kovács, P., Kutassi-Kovács, S., and Kajtár, J. (1984) GPC, CD and sedimentation analysis of poly-Lys and branched chain poly-Lys-poly-DL-Ala polypeptides. *Colloid Polym. Sci.* **262,** 208–212.
20. Mezö, G., de Oliveira, E., Krikorian, D., et al. (2003) Synthesis and comparision of antibody recognition of conjugates containing herpes simplex virus type 1 glycoprotein D epitope VII. *Bioconjugate Chemistry* **14,** 1260–1269.
21. Van der Ploeg, J. R., Drijfhout, J. W., Feijlbrief, M., Bloemhoff, W., Welling, G. W., and Welling-Wester, S. (1989) Immunological properties of multiple repeats of a linear epitope of herpes simplex virus type 1 glycoprotein D. *J. Immunol. Methods* **124,** 211–217.
22. Carfi, A., Willis, S. H., Whitbeck, J. C., et al. (2001) Herpes simplex virus glycoprotein D bound to the human receptor HveA. *Molecular Cell* **8,** 169–179.
23. Carlsson, J., Drevin, H., and Axen, R. (1978) Protein thiolation and reversible protein-protein conjugation. N-Succinimidyl 3-(2-pyridyldithio) propionate, a new heterobifunctional reagent. *Biochem. J.* **173,** 723–728.
24. Matsueda, R. and Walter, R. (1980) 3-nitro-2-piridinesulfenyl (Npys) group. *Int. J. Peptide Protein Res.* **16,** 392–401.
25. Mezö, G., Mihala, N., Andreu, D., and Hudecz, F. (2000) Conjugation of epitope peptides to branched chain polypeptides via Cys(Npys). *Bioconjugate Chemistry* **11,** 484–491.
26. Hudecz, F., Hilbert, Á., Mezö, G., et al. (1993) Epitope mapping of 273-284 region of HSV glycoprotein D by synthetic branched polypeptide carrier conjugates. *Peptide Res.* **6,** 263–271.
27. Stuchbury, T., Shipton, M., Norris, R., et al. (1975) A reporter group delivery system with both absolute and selevtive specificity for thiol groups and an improved fluorescent probe containing the 7-nitrobenzo-2-oxa-1,3-diazole moiety. *Biochem. J.* **151,** 417–432.
28. Kaiser, E., Colescott, R. L., Bossinger, C. D., and Cook, P. I. (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34,** 595–598.

III

Practical Guides

14

Protein Identification by Mass Spectrometric Analyses of Peptides

Ashley Martin

Summary

The focus of this chapter is the isolation and analysis of subsets of proteins or subproteomes. These methodologies can be applied to address fundamental questions in many areas of biological research. Protocols are provided for common techniques that include protein isolation from mammalian cells and the subsequent identification of peptides derived from endogenous proteins by the action of proteases.

Key Words: Mass spectrometry; proteome; SDS-PAGE; trypsin digest; cell lysis.

1. Introduction

The field of proteomic research is now very large and diverse and as such it is not possible to cover all common techniques. This article will, therefore, focus on the isolation and analysis of subsets of proteins. These subproteomes are intended to select for proteins that are related in some biologically meaningful way—for example, proteins that interact with each other or proteins that contain a carbohydrate moiety or a phosphate group. These methodologies are commonly utilized to address the sorts of questions that are asked in many areas of biological research. Therefore, this chapter will not consider some very interesting techniques such as "shotgun" proteomics or ICAT technology ***(1)*** that address the analysis of global or very complex proteomes using multidimensional chromatography protocols. The reader should be aware that the methods given here are those that we have found to be useful in our work, but alternative approaches and strategies are available and thought should be given to the methods adopted.

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

One common method that will not be considered here is two-dimensional electrophoresis of proteins prior to protein identification. This topic has been extensively reviewed elsewhere ***(2)***. Furthermore, this chapter will not attempt to review the different instruments available to analyze the samples using mass spectrometry. Some instruments are capable of analyzing proteins directly (for example, MS analysis using a MALDI-based platform) and others analyze peptides (for example, LC MS/MS), derived from endogenous proteins by the action of proteases. The protease employed is very often, but not exclusively, trypsin and the analysis of the peptides allows the parent protein to be identified.

2. Materials

1. Cell washing solution: This should be an isotonic, protein-free solution at a neutral pH, such as phosphate-buffered saline or Hank's buffered salt solution.
2. Lysis solution: The composition of the lysis solution varies in each system under study.
3. Antibody: this should be the highest quality available.
4. Antibody capture system, often a 4% agarose bead (Sepharose) conjugated to protein A or protein G.
5. 0.02% Sodium thiosulfate (0.1 g $Na_2S_2O_3$ in 500 mL distilled water, prepared fresh).
6. 0.1% Silver nitrate (0.5 g $AgNO_3$ in 500 mL distilled water—store in the dark).
7. Developer: 10 g Na_2CO_3, 540 µL formalin 37% in 500 mL distilled water—store at room temperature, replace when it turns yellow.
8. 0.08% Coomassie brilliant blue G250 (CBB 250)/1.6% orthophosphoric acid/8% ammonium sulfate/20% methanol.
9. Solvents: these should all be of the highest quality available (*see* **Note 1**).
10. Trypsin: from Promega (Madison, WI) (V5111 modified sequencing grade).

3. Methods

Some of the techniques given here are very common protocols and many laboratories will have standard conditions, a number of which will be suitable for this work. The methods given describe how to (1) prepare samples, (2) isolate subproteomes, and (3) process the samples to produce peptides for MS analysis.

3.1. Sample Preparation

The sensitivity and quality of data available from modern mass spectrometers mean that the preparation of samples is one of the main limitations in many proteomic protocols.

3.1.1. Cell Lysis

For cells in culture the dish is washed with ice-cold wash solution three times to ensure that all of the protein in the tissue culture medium is removed. Care should be taken to remove all the wash solution in the final wash, after which lysis

solution is added to the dish. The volume of lysis solution is best kept to a minimum, so just a sufficient amount to cover the cells should be used. Approximately 1 mL is enough to harvest a standard T75 tissue culture flask. The cells are then incubated at 4°C for 30 min with constant rocking to ensure complete lysis of the cells, which are then scraped off the dish. If required several methods can be employed to ensure complete solubilization of the proteins, such as sonication, passing the sample through a fine-gage needle, or homogenization, although care should be taken not to disrupt the protein–protein interactions. The lysate is then spun in a microfuge for 10 min at 10,000*g* to pellet any insoluble material ***(3)***.

3.1.2. Tissue Lysis

Tissues, in general, present problems because they are held together by connective materials as part of their physiological architecture, making disruption of the material more difficult. Specific problems can exist for some tissues—for example, breast and adipose tissues are often very fatty or muscle can be difficult to homogenize because of the fibrous connective material—so tissue-specific protocols may need to be developed to ensure that a good lysate is obtained. If possible the tissue should be perfused with an isotonic solution to remove blood, which should also be rinsed off the surface of the tissue. Care must also be taken to ensure that the proteins are extracted efficiently. One potential difficulty, even with small pieces of tissue, is that only cells in the outer sections of "lumps" will be completely disrupted by the lysis solution. For some tissues a conventional homogenization can be performed and for those tissues that do not homogenize easily it is possible to grind the tissue into a fine powder, in liquid nitrogen with a mortar and pestle, allowing a subsequent complete and reproducible lysis. The lysate is then centrifuged to remove any insoluble material.

3.1.3. Composition of Lysis Buffer

The composition of the lysis solution is dictated by the nature of the proteins under study and the subsequent techniques applied to the sample. One of the major choices to be made is whether or not a detergent is required at this stage. If the membrane and soluble fractions are to be separated the initial cell disruption protocol should not include a detergent, as many of the membrane proteins would be solubilized. In this case physical disruption of the cells should be used (e.g., sonication of cells or homogenization of tissues). The choice of lysis conditions is a vital consideration in this work, as proteins need to be solubilized while preservation of posttranslational modifications, inhibition of proteases, maintenance of protein–protein interactions, and, if an immunoaffinity purification step is to be performed, suitability for the antibody to function are essential. For example, SDS is very good at solubilizing membrane proteins but

most antibodies (often used to isolate proteins) do not function in the presence of this detergent (at the concentrations required to solubilize proteins) and protein–protein interactions would probably be disrupted (*see* **Note 2**). The MS analysis protocol also needs to be considered, as some detergents can cause problems with the analysis (*see* **Note 3**). Therefore, the considerations include the use of phosphatase inhibitors, protease inhibitors, pH and salt type and concentration, in addition to choice and concentration of detergent.

3.2. Isolation of Proteins

A relatively common question addressed by proteomics is the identification of proteins that interact with a target of interest to produce a protein complex. If an antibody to one member of the complex is available it is theoretically possible to isolate and identify all the associated proteins. Alternatively, an epitope tagged (for example, FLAG or HA) version of a protein can be expressed in a cell and the complex isolated by use of an antibody raised against the tag.

The isolation protocol is essentially an immunoprecipitation or immunoaffinity purification where an antibody to a defined target is added to a sample and allowed to bind to the target. The antibody–target complex, along with any associated proteins, is then isolated and the complex analyzed. The complex is often isolated using a Sepharose bead that has protein A or G covalently attached to it. This type of bead is in very common usage and the choice of which to use is dictated by the antibody source and type ***(3)***. Several variations of this basic protocol exist and many workers will have optimized protocols for their own systems. It should also be remembered that conditions developed to optimize binding of the epitope containing target protein may not be suitable for this proteomic work as the interacting proteins are the ones of interest. A basic overall protocol is given here but many variations are possible to improve the quality of the samples, some of which are discussed in **Notes 3** and **4**.

3.2.1. Preparation of Beads

The antibody capture bead (for example, protein G) is washed three times with lysis buffer by spinning the beads in a microfuge at 4000*g* for 2 min. The supernatant is aspirated and replaced with lysis solution, the beads resuspended, centrifuged again, and the whole process repeated at least twice. At the end of the process the beads are suspended with lysis buffer in a volume approximately equal to the volume of packed beads.

3.2.2. Pre-Clearing Lysate

50 µL of the bead/lysis buffer mixture is added to each lysate, which is then incubated for 1 h at 4°C with constant mixing. The suspension is then centrifuged at 4000*g* for 2 min and the supernatant collected using a pipet. This

process is termed "pre-clearing" and is intended to deplete the lysate of proteins that bind directly to the bead. Pre-clearing can be performed several times to improve removal of these proteins. If multiple rounds are performed, Sepharose (4% agarose) beads can be used for most of them and protein G beads just once to decrease the cost of the procedure.

3.2.3. Addition of Antibody

Once the pre-clearing is completed, enough antibody is added to the sample to bind the target protein and the mixture incubated for 1 h at 4°C with constant mixing. Some systems may utilize significantly different conditions from those given here; for example, an overnight incubation with the antibody is often employed. This can be used but while the longer incubation may result in a greater recovery of the target protein, some of the associated proteins may be lost.

3.2.4. Harvesting Antibody Complex

Sufficient bead suspension to bind all the antibody is then added to each sample and another 1 h incubation performed. The amount of antibody and bead suspension required should be determined empirically and the minimum of each employed. The beads, which now have the antibody bound to them, are harvested by centrifugation for 2 min at 4000*g* in a microfuge.

An alternative strategy that is often employed is to pre-incubate the antibody with the bead to coat the bead with antibody. After the bead is washed with lysis buffer to remove any unbound antibody, the bead–antibody complex is added to the pre-cleared lysate to bind the target protein. After 1 h incubation the beads are harvested and processed. The advantage of doing this here is that the binding of the antibody to the epitope and the subsequent isolation of the complex is more rapid than when the additions are performed separately, potentially increasing the recovery of the associated proteins.

3.2.5. Washing Beads

The sample is aspirated and the beads washed (gentle resuspension and incubated at 4°C for 15 min with occasional agitation) with 1 mL of lysis buffer a minimum of three times. The isolated proteins are now ready to analyze. Potential approaches to optimize this protocol for the isolation of interacting proteins are discussed in **Notes 3** and **4**.

3.3. Sample Processing

The earlier sample isolation protocol (*see* **Subheading 3.2.**) produces bead-associated proteins that are processed following dissociation of the proteins from the bead. Proteins released intact can be analyzed directly using a MALDI-based mass spectrometer, although this would be very unlikely to identify the pro-

teins in the sample and would normally be of limited use. The more usual approach is to digest the proteins using a protease (most often trypsin) to generate peptides that are subsequently analyzed. The protease treatment can be performed directly on the released proteins or after separation using SDS-PAGE.

3.3.1. Release of Proteins From Bead

The bound proteins can be liberated from the bead by incubation with 100 m*M* glycine pH 2.0 ***(3)***. This approach elutes all the proteins associated with the bead, including the antibody and contaminating proteins. It is possible to specifically elute the target protein (along with any associated proteins) if the epitope to which the antibody binds is known (for example, epitope tagged proteins) by adding synthetic peptide epitope to compete for binding to the antibody. This "affinity elution" step can be very useful, as the nonspecific contaminants that bind to the bead should not be eluted. The amount of peptide used for the elution should be determined for each system and the minimum concentration that gives effective elution employed, because too much peptide may nonspecifically displace proteins. An analogous approach that does not rely on an antibody is to employ a His-tagging protocol and to bind the protein to a nickel-loaded metal chelate column followed by elution with imidazole or EDTA.

3.3.2. Preparation of Proteins

It is possible to digest the proteins eluted from the beads directly, but it is more usual to separate the proteins using SDS-PAGE prior to the analysis. The reasons why it is normally preferable to separate the proteins are given in **Note 5**. SDS-PAGE is a very standard methodology that will not be discussed in detail here, but two specific points related to MS analysis should be considered. Methionine residues are often oxidized with this methodology, increasing the mass of the residue by approx 16 Da. Cysteine can also be modified by free acrylamide monomers producing the acrylamide adduct propionamide, increasing the molecular weight of the residue by approx 71. The database search performed at the end of the analysis can take such potential modifications into account ***(4)***.

3.3.3. Staining of Proteins

The choice of stain used will depend on the experiment being performed. Here, two of the more commonly employed MS compatible stains will be given, although others are available.

3.3.3.1. Silver Stain

This method is taken from the work of Shevchenko et al. ***(6)***. All steps are performed at room temperature with gentle shaking on an orbital platform shaker. A clean tray that is slightly larger than the gel should be used.

1. Fixation: Once the electrophoresis is complete the gel is incubated in a mixture of methanol/acetic acid/distilled water (45/5/45) for a minimum of 30 min but can be left overnight if required.
2. Wash: The fixing solution is gently poured off and replaced with distilled water and incubated for 20 to 60 min. The wash is repeated twice more to remove all the acetic acid.
3. Sensitizer: The gel is treated with fresh 0.02% sodium thiosulfate for 2 min.
4. Wash: Two incubations with water for 1 min each to remove the excess sodium thiosulfate.
5. Stain: The washed gel is then incubated with 0.1% silver nitrate for approx 30 min.
6. Wash: Two incubations with water for 1 min each to remove the excess silver nitrate.
7. Developer: The gel is then incubated with 0.04% formaldehyde/2% sodium carbonate in water to develop the color. The gel should be watched closely as the stain develops, which can take from 30 s to 3 min.
8. Quench: Once the staining has become deep enough the staining solution is rapidly poured off and replaced with 1% acetic acid to stop the reaction.
9. Storage: The stained gels can be stored in the 1% acetic acid solution at 4°C until the bands are analyzed. This is often done by sealing the gel in a plastic pouch.

3.3.3.2. Colloidal Coomassie Stain

A version of this method can be found in **ref. *3***.

The staining solution should be made up by dissolving 40 g of ammonium sulfate in 384 mL distilled water and 8 mL orthophosphoric acid.

5% Coomassie brilliant blue G250 is prepared by adding 0.5 g of dye to 10 mL of distilled water and mixing thoroughly. This mixture should not be filtered, as the dye is not completely dissolved under these conditions.

8 mL of this dye solution is added to the ammonium sulfate/phosphoric acid. This mixture can be stored at room temperature for weeks to months until it is used following the addition, with constant mixing, of 1 part methanol to 4 parts of dye mixture. The overall composition of the dye is:

- 0.08% Coomassie brilliant blue G250
- 1.6% Orthophosphoric acid
- 8% Ammonium sulfate
- 20% Methanol

1. *Stain*: After the gels are run they are placed directly in freshly prepared stain, which also fixes the proteins. The gels are placed on a shaking platform in a closed box at room temperature and left to stain overnight.
2. *Destain*: After the stain is poured off the gel is rinsed in water twice to remove excess stain and then placed into 1% acetic acid. This should be changed regularly to speed up the destaining. Small pieces of sponge can also be added to the destain to speed up the process. The gel background will become completely clear eventually, while the proteins remain blue.

3. *Storage*: The gels can be stored in the destain solution at 4°C until the bands are analyzed. This is often done by sealing the gel in a plastic pouch.

3.3.4. In-Gel Digestion

One of the most significant problems with this work is contamination of the samples with environmental proteins (such as keratin from skin and hair cells). Therefore, care should be taken to work with clean materials and, if possible, operate in "positive pressure" hoods to decrease the potential for contamination. Also, the Eppendorf-type tubes used in these procedures should be rinsed in 0.1% formic acid dissolved in 50:50 water:acetonitrile, dried under vacuum, and stored under dust-free conditions.

The stained gel should be placed on a glass plate (cleaned using methanol just prior to use) and the bands of interest excised using a clean, sterile scalpel blade. A new disposable scalpel blade can be used for each band or the blade can be rinsed in methanol between excisions. The smallest piece of gel possible should be taken. The gel can be cut into small pieces before proceeding but this can result in small pieces of gel breaking off and causing blockages in the subsequent HPLC analysis. All the volumes quoted are based on "average" bands from mini-gels; if larger gel pieces are used the volumes can be increased.

With silver-stained gels the stain needs to be removed at this point as follows:

1. *Water wash*: The band is placed in one of the pre-washed Eppendorf tubes and 300 µL of distilled/deionized water added. After 15 min the water is removed carefully using a pipet with a fine tip. This step removes the acetic acid contained in the destain solution used for storage.
2. *Silver removal*: 200 µL of 15 m*M* potassium ferricyanide/50 m*M* sodium thiosulfate is added and the tube incubated, with shaking to aid diffusion, until the band is clear. This can take between one and several hours. Once complete the supernatant is removed and the gel washed with 300 µL of water.

   After this step the silver- and Coomassie-stained gels are processed identically.
3. *Wash and pH equilibration*: 100 µL of 50 m*M* ammonium bicarbonate in 50% acetonitrile is added to the tube. This is then incubated, with shaking, for 45 min at 37°C. This incubation adjusts the gel to pH 8.0 and removes any residual contaminants. The acetonitrile helps to maintain the fixation of the protein to decrease any potential losses.

   The previous step is repeated to ensure complete equilibration. After the supernatant is removed the gel is dried down using a "speed vap." This removes the liquid from the gel so that when solutions are added back to the dried gel the constituents of the solution enter the gel along with the solute.
4. *Reduction*: 50 µL of 5 m*M* DTT (made up fresh in 100 m*M* ammonium bicarbonate in 10% acetonitrile) is added and an incubation of 1 h at 60°C performed, after which the supernatant is removed. This step should help to denature proteins (effect

of the temperature and reduction of disulfide bonds), which will aid the subsequent proteolysis.

5. *Carboxymethylation*: 50 µL of 10 m*M* iodoacetamide (made up fresh in 100 m*M* ammonium bicarbonate in 10% acetonitrile) is added and the tube incubated for 30 min at room temperature in the dark. The supernatant is removed and the gels washed as follows. This step is performed to carboxymethylate the cysteine residues to ensure that disulfide bonds cannot re-form. This increases the molecular weight of the residue by approx 57 Da and needs to be taken into account when the database search is performed ***(4).*** Iodoacetic acid can be employed, rather than iodoacetamide, but this causes an increase of approx 58 Da in the weight of cysteine residues.
6. *Wash*: 100 µL of 40 m*M* ammonium bicarbonate in 10% acetonitrile is added and incubated for 15 min at room temperature, with shaking, and then the supernatant removed. This wash is performed three times before the gel is completely dried using a "speed vap."
7. *Trypsinization*: The washed, dried gel is rehydrated using 20 µL of 12.5 µg/mL Promega sequencing-grade modified trypsin dissolved in 40 m*M* ammonium bicarbonate/10% acetonitrile for 1 h at room temperature. As this solution rehydrates the dried gel, the trypsin enters the gel. If all the liquid is soaked up and the gel still has some dry areas, more of the trypsin solution should be added and the gel allowed to swell again. Sufficient solution should be added to ensure that the gel completely rehydrates using this trypsin-containing solution. A small excess of this solution does not cause any serious problems but an insufficient amount will result in incomplete digestion of proteins. This rehydration process allows the trypsin to enter the gel and to come into intimate contact with the fixed proteins, which would not happen if the gel were not dried prior to the addition. After the gel is completely re-swelled, 20 µL of 40 m*M* ammonium bicarbonate in 10% acetonitrile is added and an overnight incubation at 37°C performed.
8. *Harvesting tryptic peptides*: The supernatant is carefully removed, replaced with 30 µL 3% formic acid in water, and incubated for 1 h at 37°C, and this solution collected and added to the initial supernatant. Another 30 µL 3% formic acid in water is added and after a further 1 h incubation the final supernatant is harvested and pooled with the two others. The approximate volume of the combined supernatants should be 80 µL. This solution is suitable for direct analysis using an LC MS/MS based system but some platforms may require a lower sample volume depending on the analytical approach. The MS analysis is considered next.

3.4. Mass Spectrometric Analysis

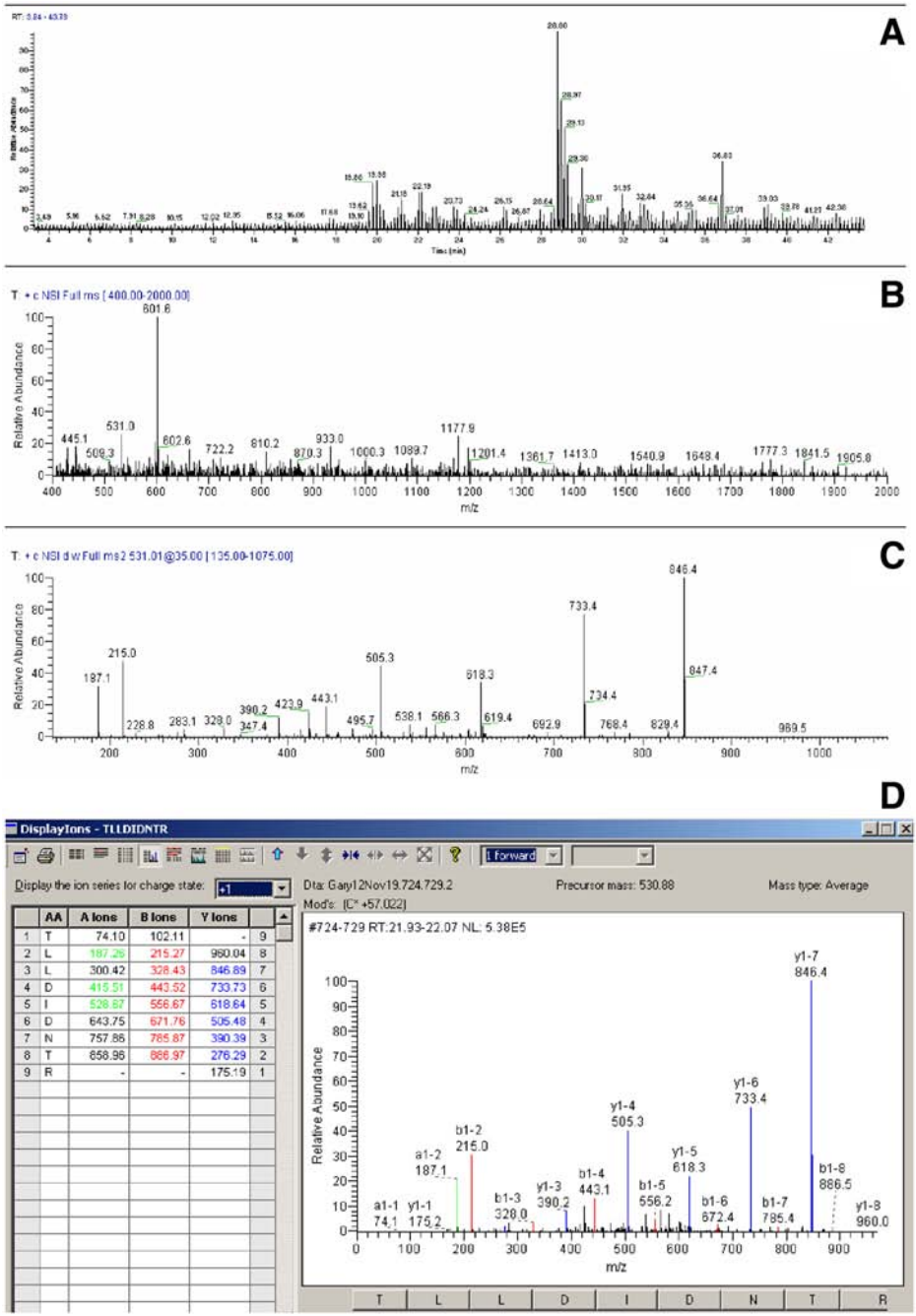
The methods given in **Subheading 3.3.** for sample preparation can be used to prepare samples for analysis using a variety of MS platforms. For illustrative purposes I will consider the analysis using a ThermoFinnigan LCQ Deca XP Plus Ion-Trap linked directly to an LC Packings/Dionex Ultimate nanobore HPLC system. Many other systems are suitable for the analysis of tryptic peptides; the

major constraint in the choice of analytical platform is likely to be the availability of access to an instrument or the cost of purchasing one. A specimen set of data is shown in **Fig. 1**, where a protein was separated using SDS-PAGE, stained, excised, and digested. The details of the analysis are given in the figure legend.

In our laboratory we have equipped the Famos autosampler (LC Packings, Amsterdam) so that samples of up to 85 µL can be analysed. This means that the samples do not need to be dried down to decrease the volume of sample before loading the peptides onto the reverse phase column. The drying-down step results in significant losses of peptides and should be avoided if possible.

The details of the analysis shown in **Fig. 1** are specific to that particular system but more generally instruments capable of doing an MS/MS analysis carry out the same basic procedure of performing an initial MS analysis to select ions for the subsequent MS/MS analysis. The selected ions are then subjected to collision-induced disassociation and the mass of the fragments determined. The fragmentation pattern and the mass of the parent peptide are then used to search various databases using a variety of search engines to identify the peptide sequence and to assign the peptide to a particular protein. Some of these programs, for example TurboSequest, are only available if the specific instrument is purchased; others are freely available via the World Wide Web.

Fig. 1. *(Opposite page)* The peptides were prepared using the in-gel digestion protocol and then separated onto a 75-µm i.d. 15-cm-long C18 PepMap reverse phase column (3 µm particle size) (Dionex) using a 5%–37.5% solvent B gradient over 40 min (solvent A = 5% acetonitrile/0.1% formic acid and solvent B = 95% acetonitrile/0.1% formic acid) at 200 nL/min using a Dionex/LC Packings Famos/Switchos/Ultimate nanobore HPLC system. The separated peptides were sprayed, using electrospray ionization, directly into a ThermoFinnigan LCQ Deca XP Plus ion-trap mass spectrometer via a nanospray source equipped with a 10-µ*M* uncoated emitter tip. A charge of 1.3 KV was applied directly to the liquid stream and the mass of the peptide determined in the initial MS analysis; these data were used to select ions for MS/MS analysis. The most abundant 5 ions for each full scan were analyzed sequentially before another full scan was performed. This data-dependent process selected ions for analysis twice, after which they were placed on an exclusion list for 1 min. **Fig. 1A** shows the base peak data against time and **Fig. 1B** shows 1 MS profile of the approx 2000 generated during the analysis. The retention time was 22.02 min. **Fig. 1C** shows the MS/MS analysis using 35% collision energy to induce fragmentation of the peak of mass 531.0 ($M+H^+$). The frag-mentation spectra were analysed using TurboSequest as part of the BioWorks 3.1 suite of data analysis programs, **Fig. 1D**. The a, b, and y series of fragments actually identified in the spectra are shown in green, red, and blue, respectively. The peptide sequence was identified as TLLDIDNTR with a cross-correlation score of 2.5.



	AA	A Ions	B Ions	Y Ions	
1	T	74.10	102.11	-	9
2	L	187.26	215.27	960.04	8
3	L	300.42	328.43	846.89	7
4	D	415.51	443.52	733.73	6
5	I	528.67	556.67	618.64	5
6	D	643.75	671.76	505.48	4
7	N	757.86	785.87	390.39	3
8	T	858.95	886.97	276.29	2
9	R	-	-	175.19	1

4. Notes

1. Solvents: The quality of solvents used in the MS analysis is very important, as contaminants can be detected during the analysis and cause severe problems. It is unusual for manufacturers to perform an MS analysis of solvents and it is advisable to try several sources of acetonitrile, water, and particularly formic acid to find suitable reagents.
2. Lysis conditions: In order to outline how the lysis conditions can be optimized for the isolation of subproteomes, immunoaffinity purification of protein complexes will be considered. The same sorts of considerations also apply to the isolation of His-tagged proteins using a nickel-loaded metal chelate resin and glycoproteins using lectins (concanavalin A or wheat germ agglutinin) bound to Sepharose beads. The optimization does not utilize MS analysis but relies on detection of the proteins using a sensitive protein stain of SDS-PAGE separated proteins.

   The screen is performed by lysing the cells using a range of conditions (detergent type/concentrations, various buffers/pHs, and salt concentrations) in the presence of protease and phosphatase inhibitors. The target protein is then immunoprecipitated employing consistent conditions and the antibody and associated proteins harvested using the relevant capture system. The beads are then washed three times using the same solution with which the cells were originally lysed. The entire protein content should then be analyzed by boiling the washed beads in SDS sample buffer and the bead/sample buffer mixture loaded onto a gel for the proteins to be separated using SDS-PAGE and the proteins detected using a sensitive silver stain. Alternatively, the target protein can be affinity eluted (if a suitable system is available) prior to the analysis to decrease the number of contaminating proteins.

   The controls are important in this work and should include an antibody-free condition to allow for proteins that bind directly to the bead and a lysate-free sample to account for the protein capture system employed. The optimum conditions can be considered to be those that give the maximum number of specific (found in the test but not control) bands. Consideration should be given to the number of nonspecific proteins, as these may make it difficult to identify the specific ones.
3. Detergent choice: When studies to optimize the protocol are performed it should be remembered that the purpose of the work is to identify proteins that interact with the target. As outlined above an initial consideration should be the choice of lysis buffer to ensure that protein–protein interactions are maintained while effective solubilization is achieved. Therefore, it may be necessary to make compromises to maximize the results. For example, extensive washing of the beads may remove contaminating proteins, but may also elute some of the interacting proteins. Similarly the choice of detergent is vital, as some detergents may help to remove contaminating proteins but again may elute the interacting proteins. The tetraspanin membrane proteins have been used to assess the effect of different detergents on protein–protein interactions. This work has shown that detergents such as CHAPS and Brij98 are very good at maintaining tetraspanin protein–protein interactions, while Triton X-100 and digitonin disrupt many nonimmunogenic

interactions ***(6)***. It is not clear if these results can be applied generally, but it does show that care should be taken when optimizing the system. Another potential problem with choice of detergent is related to problems seen in the MS analysis. Detergents such as Nonidet P40, Triton X-100, and Brij98 are termed "polyethylene" detergents. This means that the detergent is not a single molecular species but is a mixture of molecules that contain different numbers of ethylene units that can be detected as a series with an increasing mass of approx 44 Da. This can cause severe problems in the MS analysis of peptides as the detergents ionize very efficiently, leading to suppression of the peptide ionisation. This does not prohibit the use of these detergents in the lysis buffer if the detergent is removed during the protocol, for example during the SDS-PAGE/in-gel digestion.

4. Decreasing contaminants: When performing immunoaffinity purification for proteomic analysis, one of the most significant problems is contamination with nonspecific proteins. Many of these appear to be related to proteins binding to the bead used to harvest the antibody–protein complex. The "pre-clearing" step given in **Subheading 3.2.2.** may help to decrease this problem, but is unlikely to cure it.

   The affinity elution steps possible—for example, the peptide epitope when using an antibody, imidazole or EDTA when using a His-tag, or specific sugar residue when using a lectin column to purify glycoproteins—are likely to improve the specificity of the isolation process.

   The amount of antibody required to bind the target protein in a system should be empirically determined and then the amount bead required to bind the antibody determined. Very often far more antibody than needed is used and a vast excess of bead is employed because conditions developed for immunoprecipitations are employed. It is a mistake to use such conditions when performing a proteomic analysis, as the contaminants do cause serious problems and the time taken to optimize the protocol is time well spent.

5. Advantages of separating proteins prior to analysis: Once the bead-associated proteins are isolated the proteins are released from the bead and analyzed directly or the proteins can be separated using SDS-PAGE prior to the analysis. The choice of route to take depends on each protocol used, but in our work prior separation of the proteins has proven to be very useful for several reasons, some of which are provided.

   Once separated and the proteins are fixed and stained, they are preserved very effectively under these conditions, allowing storage for long periods of time prior to analysis. SDS-PAGE gives an indication of the molecular weight of the proteins under investigation, which can help when assigning protein identifications. If the appropriate antibody-free controls are performed, separation and staining of the proteins allows the contaminating proteins to be identified, decreasing the possibility that false-positives will be detected.

   The overall SDS-PAGE/in-gel digestion protocol is useful as it removes many potential contaminants; for example, salts are removed and detergents such as NP40 and TX-100 are displaced by the SDS and removed from the sample. The SDS is subsequently lost during the in-gel digestion protocol given above, producing a

very clean sample. Furthermore, if a peptide epitope was employed to release the target protein from the antibody the peptide could potentially cause problems in the subsequent MS analysis and SDS-PAGE would separate the proteins of interest from the peptide.

References

1. Patterson, S. D. and Aebersold, R. H. (2003) Proteomics: the first decade and beyond. *Nat. Genet. (suppl.)* **33,** 311–323.
2. Herbert, B. R., Harry, J. L., Packer, N. H., Gooley, A. A., Pedersen, S. K., and Williams, K. L. (2001) What place for polyacrylamide in proteomics? *Trends Biotechnol.* **19,** S3–S9 Suppl. S.
3. Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Press, Cold Spring Harbor, NY.
4. ABRF Delta Mass reference database. Web Page http://www.abrf.org/index.cfm/dm.home
5. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Analyt. Chem.* **68,** 850–858.
6. Serru, V., Le Naour, F., Billard, M., et al. (1999) Selective tetraspan-integrin complexes (CD81/α4β1, CD151/α3β1, CD151/α6β1) under conditions disrupting tetraspan interactions. *Biochem. J.* **340,** 103–111.

15

Manual Solid-Phase Synthesis of Glutathione Analogs

A Laboratory-Based Short Course

Ursel Soomets, Mihkel Zilmer, and Ülo Langel

Summary

This chapter provides a manual for a laboratory-based short course to introduce the common techniques of solid-phase peptide synthesis (SPPS). The course provides students the opportunity to design and manually synthesize analogs of glutathione using relatively simple equipment available in any unsophisticated laboratory. The manual provides compact protocols for both the different steps of SPPS and the final cleavage of peptides from resin supports. We also introduce a simple method for the synthesis of combinatorial libraries of glutathione analogs that is suitable for those relatively unfamiliar with the field of peptide chemistry.

Key Words: Solid-phase peptide synthesis; *t*-Boc chemistry; Fmoc chemistry; glutathione.

1. Introduction

The generation of a series of peptides for structure-activity relationship (SAR) studies is one of the most common applications of peptide synthesis. Such methods commonly entail the synthesis of a large number of peptides of an intermediate size of less than 30 amino acids. For larger peptides/small proteins a molecular biology approach requiring the expression of sequences as fusion proteins, is the primary method of choice. Thus, the methods presented here are preferable for the generation of libraries of smaller peptides that could be used as ligands for peptide receptors, novel antioxidants, antigens in diagnostic kits and vaccines, novel cell-permeable carriers, and the like. An additional advantage of solid-phase peptide synthesis (SPPS), compared with fusion protein expression, is the possibility of readily introducing noncoded amino acids and peptide backbone surrogates.

From: *Methods in Molecular Biology, vol. 298: Peptide Synthesis and Applications*
Edited by: J. Howl © Humana Press Inc., Totowa, NJ

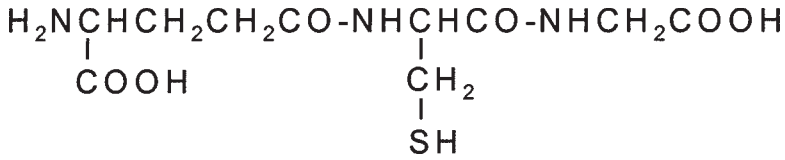


Fig. 1. Chemical structure of glutathione.

This course begins with the design of a novel glutathione analog and a variety of different approaches can be employed. One or two amino acids could be added to or substituted within the existing glutathione sequence. Modifications of the N- and C-termini or the peptide backbone (using *N*-methyl amino acids, for example) are also feasible. Data obtained from our existing studies of the tripeptide glutathione are, therefore, provided to facilitate the rational design of novel glutathione analogs.

1.1. Glutathione and Its Analogs—Promising Biomolecules

Thiol-containing compounds are central to many biochemical and pharmacological reactions. The response of cells to any stress involves changes in thiol content as thiols are consumed to protect cells via different mechanisms (e.g., antioxidativity, detoxification, signaling, direct modification/regulation of biomolecules). The discovery and isolation of glutathione, a thiol-containing compound, by Hopkins initiated "the story of glutathione" ranging through different scientific fields.

Glutathione, a tripeptide, primarily exists in the form of GSH (L-γ-Glu-L-Cys-Gly, MW 307; **Fig. 1)** and is the major low-molecular-weight thiol compound in animals and plants. GSH is present in millimolar concentration in mammalian cells ***(1***); the highest concentrations are found in the red blood cells, liver, pancreas, kidneys, spleen, eyes, lungs, and intestinal cells. The plasma concentration of GSH is 0.004 to 0.006 m*M*.

Molecules containing cysteine residues (sulfhydryl groups) easily participate in thiol-disulfide exchange and, therefore, the intracellular oxidation of GSH to glutathione disulfide (GSSG) is feasible. However, GSSG is normally maintained at less than 1% of total glutathione ***(2)*** as it is reduced to GSH. Consequently, GSH, GSSG and mixed disulfides (proteinSSG) are typical representatives of the "glutathione world" in mammalian cells, with glutathione (GSH) as the central player.

1.2. Biofunctionality

GSH is known to be a major antioxidant in human body cells that serves as a bioprotector against cell damage. Because of its ingenious structure, GSH also fulfills many other biofunctions. The isopeptidic nature of the γ-glutamyl link-

age renders GSH resistant to cleavage by most peptidases. These biochemical properties serve as the basis for the following crucial biofunctions of GSH ***(2–5)***:

- GSH prevents the denaturation of hemoglobin and reduces methemoglobin back to hemoglobin in the red blood cells;
- GSH, as the major cellular nonenzymatic antioxidant, eliminates reactive oxygen species (ROS), including the hydroxyl radical, peroxynitrite, peroxides and N_2O_3. GSH plays a principal role in cellular defense against high-grade oxidative and nitrosative stress mainly via cooperation with Se-containing glutathione peroxidase (GPx);
- GSH is necessary for the synthesis of proteins and nucleic acids, and for the processing of leukotrienes and prostaglandins;
- GSH is used for the detoxification of several xenobiotics by glutathione-*S*-transferases (GST);
- GSH is involved in the transport and storage of nitric oxide;
- GSH is involved in transport of amino acids into the liver and kidney cells;
- GSH is involved in the regulation (glutathionylation) of action of several key-enzymes and other proteins (e.g., phosphorylase, creatinine kinase, ras) and in the restoration of the sulfhydryl groups of proteins (maintenance of enzymes and proteins in active form);
- GSH acts also as a molecular regulator of whole-cellular physiology (it participates in the suppression of apoptosis, the regulation of the hexose monophosphate shunt, signal transduction, etc.);
- GSH is required for the stabilization of cell membranes and is important in the absorption of iron and selenium.

The very wide spectrum of GSH functions, and the dramatic depletion of cellular GSH by a variety of toxicological and pathological processes, explains why procedures to achieve the repletion/maintenance of cellular GSH have had a substantial impact on contemporary pharmacological and clinical developments ***(3,6–8)***.

Administration of GSH, its precursors, or analogs (acting as antioxidants, mimetics of GSH, and so on) is an adequate approach for the repletion, maintenance, or support of cellular GSH level ***(3,7–9)***. However, the administration of GSH itself has substantial limitations that include extracellular degradation, poor penetration of the blood-brain barrier, and limited direct uptake into cells. It has been shown that the GSH analog glutathione isopropyl ester (YM 737) protects against cerebral ischemia in rats by inhibiting lipid peroxidation ***(10)***. Hence, it would be of substantial pharmacoclinical interest to design novel analogs of GSH that mimic or exceed the effectiveness of GSH itself. Such GSH-like novel compounds may have impact as additional therapeutic factors (e.g., in the case of inborn defects of GSH metabolism, high-grade oxidative stress), or as potential drug precursors.

2. Materials

2.1. Reagents

1. *t*-Butyloxycarbonyl (*t*-Boc) and fluorenylmethyloxycarbonyl (Fmoc) protected amino acids.
2. MBHA and/or PAM resins for *t*-Boc chemistry and MBHA or Wang resins for Fmoc chemistry.
3. Hydroxybenzotriazole (HOBt).
4. 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylammonium tetrafluoroborate (TBTU).
5. Kaiser test reagents:

Solution A: mix solutions a and b.

a. 200 m*M* KCN in pyridine: dilute 2 mL of KCN solution (6.5 mg in 10 mL of water) in 100 mL of pyridine. KCN is highly toxic. Unused aqueous KCN solution should be stored at 4°C in the dark.
b. Dissolve 40 g of phenol in 10 mL of absolute ethanol. Phenol can cause severe burns, and should be kept in the dark.

Solution B: ninhydrin solution.

Prepare 5% (w/v) ninhydrin solution in ethanol (500 mg of ninhydrin in 10 mL of ethanol).

6. Solvents and scavengers: dichloromethane (DCM), trifluoroacetic acid (TFA), *N*, *N*-diisopropylethylamine (DIEA), 100% ethanol, *N*,*N*-dimethylformamide (DMF), dimethylsulphoxide (DMSO), acetonitrile, trifluoromethanesulfonic acid (TFMSA), dimethylsulfide (DMS), *p*-cresol, dimercaptoethane.

2.2. Equipment

1. Reaction vessels for manual synthesis.
2. Rotary evaporator.
3. Apparatus for HF cleavage (*t*-Boc chemistry only).
4. Freeze-drier (lyophilizer).
5. HPLC system with C_{18} preparative column.
6. Mass spectrometer (MALDI TOF or ESI-MS or PDMS).

3. Methods

3.1. Summary of Laboratory Procedures

3.1.1. Anchoring of First Amino Acid

Synthetic steps for the attachment of the first amino acid may differ from those of chain elongation. The subsequent methods are a commonly used for the synthesis of peptides with C-terminal carboxylic acids or amides.

3.1.2. Stepwise Solid-Phase Peptide Synthesis

A standard protocol is provided for the synthesis of the peptide you have designed. Coupling efficiencies during manual synthesis can be quantitatively or qualitatively monitored using the Kaiser test to determine the concentration or presence of free primary amino groups. SPPS is based on the sequential addition of α-amino- and side chain-protected amino acid residues to an insoluble polymeric support. The acid-labile *t*-Boc group or base-labile Fmoc group is used for N-α-protection. The reactive side chains of amino acids are protected with different protective groups. After removal of the N-terminal protective group, the next protected amino acid is added. Several different methods (e.g., DCC, HOBt, TBTU, symmetrical anhydride) to preactivate protected amino acids will be employed to provide you with experience in their use.

3.1.3. Removal of Protective Groups and Cleavage of Peptide From Resin With Strong Acid

***t*-Boc chemistry:** The synthetic peptide will be cleaved and deprotected from protective groups in two steps. First, the DNP- (with thiophenol), benzyl- (partly) and formyl- (with "low-TFMSA" method) protective groups will be cleaved. After this step, the final cleavage of the deprotected peptide from the resin will be carried out with liquid HF.

Fmoc chemistry: After removal of the N-terminal Fmoc-protective group the synthesized peptide will be deprotected and cleaved from the handle-resin with trifluoroacetic acid (TFA). The peptide will subsequently be separated from scavengers (added during the cleavage), and lyophilized.

3.1.4. Purification by Reverse-Phase HPLC

The purity of crude peptide will be analayzed on a reverse-phase high-performance liquid chromatography (HPLC) column. The elution of the peptide is followed by absorbance recorded at 215 nm. The purification on a 10 to 20 mg scale is performed on a larger, preparative column monitored at the less sensitive wavelength of 238 nm.

3.1.5. Analysis of Purified Peptide by MALDI-TOF Mass Spectroscopy

The purified lyophilized fractions of peptides (and significant side products) are commonly analysed using a matrix-assisted time-of-flight laser desorption/ionization (MALDI-TOF) mass-spectrometer. The molecular mass spectra obtained should be analyzed and compared to the calculated mass.

3.2. Resin-Handling Procedures

3.2.1. Swelling of Resin

1. Place dry resin into a reaction vessel.

2. Add DCM (3 × bed volume).
3. Shake gently for 30 min.
4. Remove DCM.

3.2.2. Washing of Resin

1. Add solvent (3 × bed volume).
2. Agitate for 1 min.
3. Drain resin to remove solvent.

3.2.3. Drying of Resin

1. In the last washing step use methanol, hexane, or ether.
2. Remove solvent and dry resin by application of a vacuum for 10 min.
3. Dry in freeze dryer overnight.

3.3. Detailed Laboratory Manual

3.3.1. Attachment of First Amino Acid to Resin

All the reactions with the resin are carried out in a glass reaction vessel of reasonable size. In all the reaction steps the volume of liquids should be kept minimal. If volumes used are too small, the effect of the washing steps will not be sufficient, while too much solvent is wasteful. Take enough of the solvents that the resin is suspended and can move freely in the reaction vessel while shaken, and not more. The amounts of some reagents are given as equivalents (eq) as compared to the amount of the free amino groups on the resin. Use your data on the amount of free amino-groups for precise calculation.

3.3.1.1. *t*-Boc Chemistry

To obtain peptides with C-terminal free carboxylic acid: The *t*-Boc-amino acid-PAM resin (with the first amino acid already attached) is used as the solid support. In this case the coupling of the second amino acid is a conventional acylation of an amino group on the resin with an activated *t*-Boc-amino acid; do not forget to deprotect the amino acid on the resin before the coupling of the following amino acid derivative. Use 200 mg of *t*-Boc-amino acid-PAM resin.

1. Deprotect *t*-Boc-amino acid-PAM-resin (after swelling 0.5–1 h in DCM): mix for 1 min with 50% TFA/DCM (v/v), remove TFA/DCM, mix for 20 min with 50% TFA/DCM (v/v).
2. Wash with DCM, EtOH, DMF, base (DIEA), DMF, base (DIEA), 3 × DCM. The resin is now in the form where the amino groups are non-protonated and can therefore be acylated with an activated amino acid.
3. Shake base-washed amino acid-resin for 30 min with 2 eq of Boc-amino acid and 2 eq of dicyclohexyl carbodiimide (DCC).

4. Wash with DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM.

To obtain peptides with C-terminal carboxyamide: this procedure should be used when you intend to synthesize a peptide with a C-terminal amide group. In this case the coupling of the first amino acid is a conventional acylation of an amino group on the resin with an activated *t*-Boc-amino acid. Use 200 mg of MBHA resin for the synthesis.

1. Wash the MBHA-resin (after swelling) for 1 min with DMF, base (DIEA), DMF, 3 × DCM. The resin is now in a form where the amino groups are in the non-protonated form and can therefore be acylated with an activated amino acid.
2. Shake base-washed MBHA-resin for 30 min with 2 eq of *t*-Boc-amino acid and 2 eq of dicyclohexyl carbodiimide (DCC).
3. Wash with DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM.

3.3.1.2. Fmoc Chemistry

Coupling of 4-(4-hydroxymethyl-3-methoxyphenoxy)-butyric acid (HMPB, for synthesis of peptide acids*)* or *p-[(R;S)-a-[1- (9H- fluorenyl- methoxyformamido]- 2,4- dimethoxybenzyl] - phenoxyacetic* acid (modified Rink linker, for synthesis of carboxamide peptides*) linkers to MBHA resin*: For Fmoc chemistry several types of solid supports are available, which include hydroxymethyl-based, aminomethyl-based, and trityl chloride resins. We describe the use of the MBHA resin. In this case the respective linker (to achieve peptide acid or amide) is coupled to the resin and first amino acid is then coupled to the linker. Attachment of the linker to the resin is a reaction between the carboxyl-group of the linker and amino-group of the MBHA resin. Commercially available resins with linkers already attached could also be used.

1. Wash the MBHA-resin (after swelling) for 1 min with DMF, base (DIEA), DMF, 3 × DCM. The resin is now in a form where the amino groups are non-protonated and can therefore be acylated by an activated amino acid.
2. Shake MBHA-resin for 30 min with 1.5 eq of handle and 1.5 eq of TBTU and 1.5 eq of HOBt in the presence of DIEA (3 eq) in DMF.
3. Wash with DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM.

Attachment of the first amino acid to the linker to produce a peptide acid: The first amino acid is coupled to the hydroxyl group of the HMBP linker using symmetrical anhydride of the amino acid.

1. Dissolve Fmoc-amino acid (6 eq of amino acid per 1 eq of linker-resin) in DCM in a separate round-bottomed flask.
2. Add a solution of 3 eq of DCC in a minimum amount of dry DCM to the amino acid solution, cool to 0°C, stir and allow to stand for 30 min at 0°C.
3. Add the anhydride solution to the linker-resin. Dissolve DMAP (1 eq) in DMF and add this solution to the reaction mixture; agitate at room temperature for 1 h.

Note: Do not use this activation procedure for coupling of Arg (conversation of anhydride to the δ-lactam) or Gly (dipeptide formation).

4. Wash 5 times with DMF (1 min each).

To produce a carboxyamide peptide: The peptide will be linked to the modified Rink linker via an amide bond. The attachment of the first residue can be carried out under conditions for peptide bond formation (e.g., with TBTU) by using the activation procedures described in **Subheading 3.3.2.2.** (Methods A–E) of this chapter. Do not forget to deprotect the linker before coupling of the amino acid.

3.3.2. Stepwise Synthesis of Peptides With t*-Boc- or Fmoc-Amino Acids*

The attachment of the first amino acid is a special procedure in both Fmoc- and *t*-Boc-chemistry if the intended peptide should contain a C-terminal carboxyl group. However, the stepwise synthesis in both Fmoc- and *t*-Boc chemistry follows the same general outline.

The protocol for the peptide synthesis can be divided into a number of different steps.

1. Removal of the *N*-α-protective group (*t*-Boc or Fmoc). In *t*-Boc chemistry, TFA is used; in Fmoc chemistry, piperidine or morpholine is used.
2. Washing of the resin with solvents and base (DIEA) to generate a non-protonated amino group on the resin.
3. If necessary, coupling of a handle to the resin to generate the TFA-labile bond between handle and resin (Fmoc chemistry).
4. Coupling of an activated *N*-α-protected amino acid. This step is the same in both *t*-Boc and Fmoc chemistry. The extent of the coupling reaction should be monitored by the Kaiser test or any other method.

3.3.2.1. Removal of *N*-α-Protective Groups

A. Removal of the *t*-Boc-group:

1. Wash peptide-resin with 50% TFA/DCM (v/v) for 1 min. If Cys or Met are included in the sequence, then add 2% *N*-acetylcysteine (scavenger) into the same solution.
2. Shake peptide-resin with 50% TFA/DCM (v/v) for 20 min.
3. Washing scheme (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, base (DIEA), DMF, 3 × DCM.
4. Monitor qualitatively the amount of free amino-groups by using Kaiser test.

B. Removal of the Fmoc-group:

1. Wash peptide-resin with 20% piperidine/DMF for 1 min.
2. Mix peptide-resin with 20% piperidine/DMF for 15 min.
3. Washing scheme (1 min each): 5 × DMF.
4. Qualitatively assess the amount of free amino groups using the Kaiser test.

3.3.2.2. Coupling of *N*-α-Protected Amino Acids

To couple *t*-Boc- or Fmoc-protected amino acids with high yield, different methods can be used. All the methods described should be exercised at least twice. For routine synthesis we recommend the use of TBTU/HOBt-activation or dicyclohexylcarbodiimide (DCC)-activation if TBTU/HOBt-activation gives low yield. For *t*-Boc-Arg(Tos), *t*-Boc-Asn and *t*-Boc-Gln, the DCC-activation cannot be recommended because of the side reactions.

The washing procedure, after using different coupling methods, is the same as described after TBTU/HOBT coupling.

Some general remarks:

- If the peptide chain is shorter than 10, use a coupling solvent mixture of 20% DMF in DCM; if longer than 10, use 30–50% DMF in DCM.
- Use a quantitative Kaiser test after the first and last deprotections.
- Otherwise, monitor the amount of the amino groups qualitatively after each deprotection and coupling step. A positive test produces a blue color (Pro, Asp, Asn—brown).
- If the coupling is incomplete (blue or pink color in Kaiser test), recouple or block remaining amines by acetylation. For acetylation add to the reaction vessel acetic anhydride and DIEA (1/1 v/v) in DCM and DMF (excess), mix for 5 min, wash as after coupling of amino acid.
- Both *t*-Boc-amino acids and Fmoc-amino acids can be coupled with the following methods if not indicated otherwise.

A. Activation with TBTU

1. Dissolve 1.5 eq TBTU, 1.5 eq HOBt, and 1.5 eq of the amino acid in a minimum amount of DMF in separate test tubes.
2. Add 3 eq DIEA to the solution and mix.
3. Add the solution to the resin.
4. Couple Arg, Ile, and Val for 30 min, all other amino acids for 10 min, agitating vessel gently.
5. Washing scheme for *t*-Boc chemistry (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM. Washing scheme for Fmoc-chemistry (1 min each): 5 × DMF.

B. Activation with DCC

1. Dissolve 2 eq of amino acid in DCM, add 2 eq of dicyclohexylcarbodiimide (DCC).
2. Add solution to *t*-Boc- or Fmoc-deprotected peptide-resin.
3. Add the same volume of DMF after 10 min.
4. Agitate for 30 min (45 min, if peptide chain is longer than 10 amino acids).This method is not suitable for coupling of *t*-Boc-Asn, *t*-Boc-Gln, and *t*-Boc-Arg(Tos).
5. Washing scheme for *t*-Boc chemistry (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM. Washing scheme for Fmoc chemistry (1 min each): 5 × DMF.

C. Activation with DCC/HOBt

1. Dissolve protected amino acid and HOBt (1:1; both 2 eq) in DMF in separate test-tubes.
2. Cool solution to 0°C and add 2 eq of dicyclohexylcarbodiimide (DCC), incubate for 1 h at 0°C.
3. Add solution to resin, add some DCM, and agitate for 40 min. In some protocols HOBt-activation is used as a routine (e.g., Applied Biosystems Peptide Synthesizer).
4. Washing scheme for *t*-Boc chemistry (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM. Washing scheme for Fmoc-chemistry (1 min each): 5 × DMF.

D. Using Symmetrical Anhydrides

1. Dissolve 6 eq of amino acid and 3 eq of DCC in a minimum amount of DCM separately outside of reaction vessel, and cool to 0°C.
2. Mix solutions and keep for 30 min at 0°C.
3. Add the anhydride solution to the peptide-resin.
4. Add 1 eq of base (DIEA, to neutralize the acid formed by symmetrical anhydride coupling) after 15 min.
5. Couple for 30 min at room temperature. Do not use this activation procedure for coupling of *t*-Boc-Arg(Tos), *t*-Boc-Asn, *t*-Boc-Gln and *t*-Boc-Gly.
6. Washing scheme for *t*-Boc chemistry (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM. Washing scheme for Fmoc chemistry (1 min each): 5 × DMF.

E. Using HATU Activation

1. Dissolve 5 eq of protected amino acid and 4.9 eq of HATU in minimal amount of DMF in separate test tube.
2. Add 10 eq DIEA and mix solution.
3. Add the mixture to 1 eq of the peptide-resin and agitate gently for 30 min.
4. Washing scheme for *t*-Boc chemistry (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, 3 × DCM. Washing scheme for Fmoc chemistry (1 min each): 5 × DMF.

*3.3.3. Deprotection of Peptides (*t*-Boc Chemistry)*

3.3.3.1. Removal of DNP Group From His Side Chain

Treat the peptide-resin with 20% of thiophenol solution in DMF for 1–2 h at room temperature while mixing. This procedure is essential when His(DNP) is present in a peptide sequence. In order to avoid the side products, the procedure must be applied before the final deprotection from *t*-Boc and formyl groups.

Washing scheme (1 min each): DMF, EtOH, 3 × DCM.

3.3.3.2. Removal of the Formyl Group From Trp(For)-Containing Peptides

A. Low-TFMSA Method: This method is used mainly to deprotect the peptide-resin from the formyl group on Trp and benzyl types of protective groups. In order to avoid the side products, the procedure must be applied after the

final deprotection from Boc-groups. The protective groups that are not cleaved are Cys(4-MeBzl), His(DNP) and Arg(Tos). His(Bom) is incompletely removed. Methionine *S*-sulfoxide is reduced to the methionine. Dimethyl sulfide (DMS) is the nucleophilic reagent in the system and attacks the benzylic positions of the protective groups. TFMSA is an extremely strong acid: use personal lab-protection and always work in a fume hood.

1. Mix deprotection mixture (TMFSA/TFA/DMS/*p*-cresol/dimercaptoethane; 10:50: 30:8:2 vol.%) on ice. Stir solution and add TFMSA dropwise as a last step.
2. Add 1 mL of deprotection mixture to 50 mg of the peptide-resin; shake for 2 h on ice.
3. Washing scheme (1 min each): DCM, EtOH, DMF, base (DIEA), DMF, base (DIEA), DMF, 3 × DCM.

B. Piperidine Deformylation

1. Cool 10% piperidine in DMF in an ice bath and add to the peptide-resin (10 mL/1 g)
2. Agitate at 0°C for 2 h.
3. Washing scheme (1 min each): 2 × DMF, 2 × DCM, 3 × EtOH.

3.3.4. Cleavage of Peptides From Solid Supports

3.3.4.1. *t*-Boc Chemistry

The distillation of HF should be carried out as a demonstration only. You have to prepare the reaction vessel for cleavage and separate the peptide from resin and scavengers. The subsequently described procedure is for 1 g of resin. If several peptide-resins are to be simultaneously cleaved, vessels should contain about the same amount of resin. Before the final cleavage, deprotect the amino terminal of the peptide.

1. Weigh 1.0 g of lyophilized peptide-resin into cleavage-vessel; cover it with 2 mL of *p*-cresol (if Cys or Met are present use a mixture of 1 mL *p*-cresol and 1 mL *p*-thiocresol).
2. After freezing the mixture to –70°C, add a stirrer bar.
3. Condense HF (20 mL HF/g of peptide-resin) to the reaction vessel at –70°C (use MeOH/dry ice).
4. Increase the temperature slowly to 0°C, incubate for 0.5–1 h at 0°C.
5. Aspirate HF in a vacuum at room temperature.
6. To separate peptide from resin and scavengers, extract the mixture 3 times with cold diethyl ether; centrifuge.
7. Dissolve the peptide in 20% acetonitrile/water, filter the peptide solution into a clean reaction vessel, and lyophilize.

3.3.4.2. Fmoc Chemistry

With the Fmoc approach, it is possible to use TFA, rather than a super-strong acid such as HF, to both cleave and deprotect peptides from linker-resins. The

Table 1
TFA Cleavage and Deprotection Conditions in Fmoc Chemistry

Peptide contains:	Arg (Mtr)	Met, Cys (Trt)	Trp	Cleavage conditions		
				TFA, vol. %	Scavenger(s), vol. %	Time, h
1	–	–	–	95	Water—5	1.5
1A	–	–	–	95	TIS/water—2.5:2.5	1.5–3
2	+	–	–	95	Thioanisole—5	3
3	+	–	–	95	Phenol—5	6
4	–	+	–	95	Ethylmethylsulfide—5	1.5
5	–	–	+	95	EDT/anisole—2.5:2.5	1.5
6	–	+	+	95	EDT/anisole/ethyl methyl sulfide—1:3:1	1.5
6A	–	+	+/–	94	TIS/water/EDT—1:2.5:2.5	1.5–3
7	+	+	+	93	Ethanedithiol/anisole/ethyl methylsulfide—1:3:3	3
7A	+	+	+	82.5	Thioanisole/water/phenol/ EDT—5:5:5:2.5	1.5–18

• For peptides containing Ser(*tert*-Bu), Thr(*tert*-Bu), Asp(O-*tert*-Bu), Glu(O-*tert*-Bu), and Lys(*t*-Boc), any of the cleavage conditions listed is suitable.

• For peptides containing trityl-protected side chains cleavage conditions 4, 5, 6 , 6A, 7, or 7A should be used.

• Cys(Acm) and Cys(*tert*-Bu) are stable in 95% TFA. These protecting groups must be removed after cleavage. In the case of these amino acids cleavage condition 3 should be used.

TFA cleavage procedure is relatively easy to perform and uses standard laboratory glassware. **Caution**: Perform this procedure in a fume hood. The following procedure is for 100 mg of peptide-resin and is carried out in a reaction vessel. Before the TFA cleavage, the amino terminal of the peptide must be deprotected from the N-α-Fmoc group.

1. Wash the peptide-resin with DCM and DMF and dry under reduced pressure.
2. Add 0.5 mL of scavengers according to **Table 1** to prevent reactions with side chains of the peptide.
3. Add 9.5 mL of TFA and agitate the vessel according to **Table 1**.
4. Remove the resin by filtration under reduced pressure and wash resin twice with TFA into the appropriate-sized round-bottom flask.
5. Evaporate the TFA and scavengers using a rotating evaporator.
6. Add cold ether (10-fold volume), mix, and transfer mixture into the centrifuge tube; centrifuge.

7. Carefully decant the ether, add fresh ether, suspend the peptide pellet, and centrifuge. Repeat three times.
8. Dissolve the peptide pellet in 20% acetonitrile/water and lyophilize.

3.4. Use of Synthetic Peptide Combinatorial Libraries to Produce Glutathione Analogs

Recently, molecular diversity technologies have been developed to accelerate the process of pharmaceutical lead discovery. Combinatorial syntheses are used to rapidly generate large number collections of compounds suitable for screening against a wide variety of biological targets. The molecular diversity field has mainly focused upon the preparation and screening of nucleic acid, synthetic peptide, recombinant peptide and, more recently, peptoid libraries. Resulting biopolymer libraries have yielded moderate- to high-affinity ligands for antibodies and receptors.

It is important to further develop the concept of structure-activity relationships to precisely define the structural requirements of glutathione action. Thus, this section introduces the design, synthesis, and screening of a peptide combinatorial library to obtain multiple glutathione analogs. Combinatorial libraries will be composed of mixtures of peptides (consisting of natural or noncoded amino acids) on solid support. After cleavage from the resin, the mixtures of the peptides will be screened directly in different specific assays.

You will synthesize a small peptide library based on glutathione, introducing three different substitutions by mixtures of four natural or noncoded amino acids. This synthesis creates a mixture of four peptides at the first step of the synthesis, 16 at the second step, and 64 at the third step (**Fig. 2**).

All the methods for SPPS are compatible with multiple peptide synthesis (MPS). The split synthesis approach will be used to create combinatorial libraries of glutathione analogs (**Fig. 2**). Equimolar amounts of amino acids will be used for each of three coupling steps. Excess amounts of amino acids known to couple more slowly can be used. After coupling, the resins will be mixed and split again into reaction vessels. When using four reaction vessels (theoretically yielding 64 compounds after the third step), the amount of amino groups on the resin should be not less than 5 mmol. That amount of amino groups on the resin will produce, after the synthesis of 3 amino acid peptides, 250 mg (yield 100% at each step for each reaction) of the individual peptide with an average MW = 500. Even in the case of reaction yields of only 10%, the obtained amount of each peptide (25 mg) is sufficient for purification and characterization. If components of the combinatorial library are difficult to separate, then a higher resin loading capacity could be used. We describe the methods of synthesis in detail in **Fig. 2**.

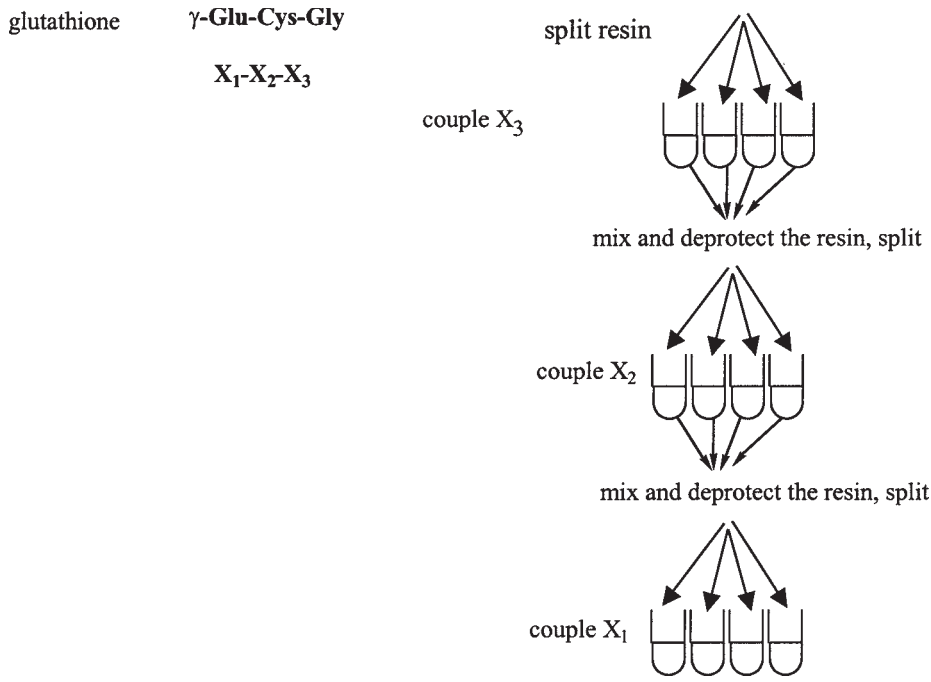


Fig. 2. Combinatorial synthesis of glutathione analogs.

The initial step in combinatorial peptide synthesis is the coupling of the first four amino acids (X_3) to the resin. After mixing of these four resins and deprotection of the mixture, the resin is split into four and a substitiuent X_2 will be coupled separately. After further mixing and splitting, the X_1 substituent is coupled. Deprotection and cleavage of peptides is performed using conventional methods (*see* **Subheading 3.3.2.1.** and **3.3.4.**) yielding mixtures of completely unprotected peptides. The mixtures of glutathione analogs can be screened for antioxidative activity.

Many different methods for the deconvolution of the components of combinatorial libraries exist today. An "active" mixture of peptides can be separated into components. Conventional reverse-phase HPLC (columns with C_{18}) has been used successfully for the separation of peptide mixtures.

3.5. Purification of Peptides by HPLC

Reverse phase HPLC (*see* **Fig. 3**) has been the main tool for both purification and analysis of peptides. Samples are eluted with a gradient of water/acetonitrile. When preparing samples for purification or analysis, materials should be completely dissolved and free from particulate matter that can rapidly block columns.

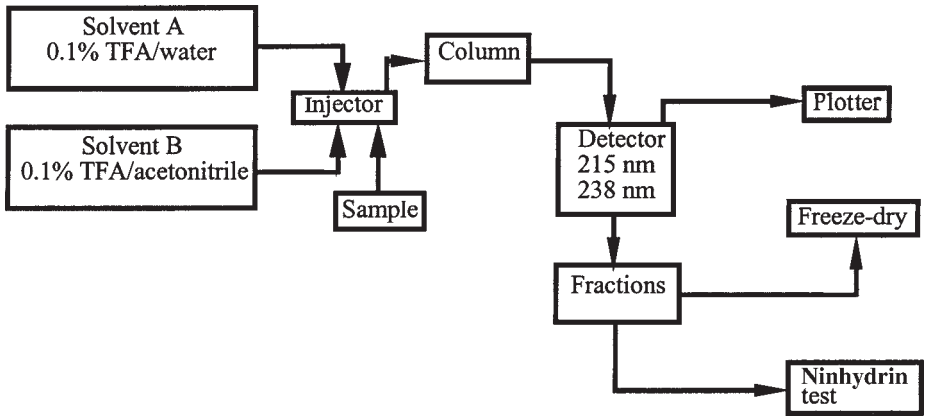


Fig. 3. A schematic representation of a RP-HPLC system suitable for the analysis and purification of peptides.

3.5.1. The Column

Stainless steel columns packed with a suitable stationary phase are used for peptide purification by HPLC. Reverse phase (RP) chromatography utilizes a stationary phase that is a nonpolar compound such as C_{18} hydrocarbon covalently bound to porous silica. For the purification of more hydrophobic peptides, C_4- or C_8-RP columns have also been used.

3.5.2. Chromatographic Solvent (Mobile Phase)

The convenient mobile phase for peptide separation is a gradient elution system, where a gradient programmer continuously changes the composition of the developing solvent. All solvents for use in HPLC systems must be of an especially pure grade both to prevent column degradation and enable the use of a highly sensitive detection system. We commonly use a gradient of 0.1% TFA/water (solvent A) and 0.1% TFA/acetonitrile (solvent B). We recommend a gradient for peptide purification from 20% B to 80% B in 50 min.

3.5.3. Detector System

For the detection of peptide fractions a variable-wavelength spectrophotometer is used. For the purpose of analysis we favor a wavelength of 215 nm that detects peptide bonds. The less-sensitive wavelength of 238 nm can be used for preparative purposes.

3.5.4. Practical Considerations

- To avoid air bubbles in solvent tubing, the solvents must be thoroughly degassed before use by stirring them under vacuum for about 30 min or in an ultrasonic bath.

- Because acidic solvents slowly destroy the reverse-phase columns, the HPLC system should be left filled with MeOH when not in use.
- To prepare 1 mL of crude peptide solution for HPLC purification weigh up 20 mg of crude peptide, add 200 mL of acetonitrile, and mix vigorously. Add 800 mL of water, mix, and centrifuge or filter to remove particles.

3.6. Analytical Tests

3.6.1. Kaiser (Ninhydrin) Test

3.6.1.1. Qualitative Test

1. Transfer a small dry resin sample (1–2 mg) after the wash period to a small test tube.
2. Add 2 to 3 drops of each Kaiser test reagent (A and B) and mix the solution.
3. Heat test tube for 5 min at 100°C.
4. If the test is negative (no free amino groups are present) beads are clear or yellow in color. A positive test (free amino groups are present) produces dark violet (dark blue)-colored beads (Pro, Asn, and Asp yield a red-brownish color).

3.6.1.2. Quantitative Test

1. Wash peptide-resin and dry carefully.
2. Weigh a test tube with a precision of 0.1 mg and add your sample to the test tube.
3. Dry the test tube under vacuum and weigh it again.
4. Add 0.7 mL of Kaiser test mixture of components A (0.5 mL) and B (0.2 mL).
5. Heat the mixture at 100°C for 10 min, dilute 200 times with EtOH, and measure optical density at 570 nm (extinction coefficient $\varepsilon = 14{,}000$).

3.6.2. Quantitative Picric Acid Test

1. Prepare reagent 1: 0.1 *M* picric acid in DCM; and reagent 2: 5% DIEA in DCM.
2. Swell the whole batch of resin in DCM (in reaction vessel).
3. Neutralize with reagent 2 (2 × 1 min); wash well with DCM (5 × 1 min).
4. Treat with reagent 1 (2 × 1 min); wash well with DCM (5 × 1 min).
5. Elute the picrate with reagent 2 (2 × 1 min); save the eluent.
6. Dilute the eluted picrate with 95% EtOH to give a suitable absorbance. The final solution should not contain more than 20% of DCM.
7. Read absorbance at 358 nm. DIEA picrate has $\varepsilon = 14{,}500$.

3.6.3. Qualitative Monitoring With Chloranil

1. Add 0.2 mL of acetone (if Pro is determined) or 0.2 mL acetaldehyde (for the other amino acids) to dried resin beads on a microscope slide.
2. Add 0.5 mL of saturated chloranil solution in toluene; mix the beads for 5 min at room temperature.
3. Blue or green resin beads indicate free amino groups.

3.6.4. Qualitative Monitoring With Bromophenol Blue (BFB)

1. Add 3 drops of 1% Bromophenol blue in DMF to the coupling reaction.
2. Intense violet color indicates the presence of free amino groups; it fades to yellow after the coupling reaction is completed.

4. Notes

1. Most of the solvents and reagents used during this course are or might be dangerous to health; therefore, special precautions should be taken in handling them. Always use gloves and eye protection.
2. The most dangerous solvents in SPPS when using *t*-Boc chemistry are hydrogen fluoride (HF) and trifluoroacetic acid (TFA). HF is extremely dangerous and requires a special HF apparatus of Teflon-coated vessels to be handled safely. TFA and TFMSA in contact with skin also cause wounds that take weeks to heal.
3. Dicyclohexylcarbodiimide (DCC) is also very reactive and can react rapidly with proteins in the skin. It can cause allergy and might be carcinogenic.
4. Dichloromethane (DCM, CH_2Cl_2) is not directly toxic unless you get heavily exposed, but as with all halogenated hydrocarbons one should always try to avoid exposure as there might be long-term effects.
5. Piperidine may cause severe irritation of skin and eye burns.

References

1. Hopkins, F. G. (1929) On gluthathione: a reinvestigation. *J. Biol. Chem.* **84,** 269–320.
1a. Meister, A. and Anderson, M. E. (1983) Glutathione. *Ann. Rev. Biochem.* **52,** 711–760.
2. Dickinson, D. A. an Forman, H. J. (2002) Cellular glutathione and thiols metabolism. *Biochem. Pharmacol.* **64,** 1019–1026.
3. Bharath, S., Hsu, M., Kaur, D., Rajagopalan, S., and Andersen, J. K. (2002) Glutathione, iron and Parkinson's disease. *Biochem. Pharmacol.* **64,** 1037–1048.
4. Filomeni, G., Rotilio, G., and Ciriolo, M. R. (2002) Cell signalling and the glutathione redox system. *Biochem. Pharmacol.* **64,** 1057–1064.
5. Huang, K. P. and Huang, F. L. (2002) Glutathionylation of proteins by glutathione disulfide S-oxide. *Biochem. Pharmacol.* **64,** 1049–1056.
6. Dringen, R. (2000) Metabolism and functions of glutathione in brain. *Progress in Neurobiology* **62,** 649–671.
7. Schulz, J. B., Lindenau, J., Seyfried, J., and Dichgans, J. (2000) Glutahione, oxidative stress and neurodegeneration. *Eur. J. Biochem.* **267,** 4904–4911.
8. Valencia, E., Marin, A., and Hardy, G. (2001) Glutathione—nutritional and pharmacological viewpoints: Part VI. *Nutrition* **18,** 291–292.
9. Lucente, G., Luisi, G., and Pinnen, F. (1998) Design and synthesis of glutathione analogues. *Il Farmaco* **53,** 721–735.
10. Yamamoto, M., Sakamoto, N., Iwai, A., et al. (1993) Protective actions of YM737, a new glutathione analog, against cerebral ischemia in rats. *Res. Commun. Chem. Pathol. Pharmacol.* **81,** 221–232.

Index

A

Aggregation, problems during SPPS, 16–17
Aib, *see* α-aminoisobutyric acid
Amino acid
  Hmb-derivative, to prevent aggregation, 16–17
  lipoamino, 46–47, 52
  one letter code, 5
  protecting group
    common, 5–7, 13–14, 157–158
    3-nitro-2-pyridinesulfenyl (Npys), 37, 83, 139, 219
  proteinogenic, 5
  side-chain, 5
  three letter code, 5
α-aminoisobutyric acid, 33
Asp-Pro sequence, decomposition in acidic solutions, 72
Automated synthesis
  instrumentation
    commercial suppliers, 176–177
    comparative applications, 174–186
  strategies, 168–174

B

Backbone amide linker (BAL), synthetic strategies, 195–208
Bioavailability, 46
Biotin–avidin interaction, 80
Blood brain barrier (BBB), 46
Bradykinin, 30, 147
Branched-chain polymeric polypeptides, 210

C

Carboxylic function activation, during peptide synthesis, 13
Cell lysis, for protein isolation, 228–229
Cell penetrating peptide (CPP)
  Antennapedia, 78
  cargoes, 79–82
  conjugation to bioactive cargoes, 80–82
  model amphiphilic peptide (MAP), 78
  penetratin (pAntp), 78
  pVEC, 78
  sequences, 78
  strategies for cargo delivery, 77–89
  Tat, 78
  toxicity, 79
  transportan, 78
Chemical ligation, for polypeptide synthesis, 113–115
Chimeric peptides
  anti-cancer agents, 33–34
  β- and T-cell epitopes, 64, 211
  cyclic, 70
  receptor ligands, 31–33
  synthesis of linear, branched and cyclic varieties, 63–76
  synthetic antigen, 64
Chimerism, as a general synthetic strategy, 25–41
Classification, of peptides, 28–29
Cleavage and deprotection
  cocktails, including scavengers, 18, 203, 251–252
  separation of peptides from resins, 18–19, 252
Combinatorial peptide libraries, 151–165, 253–254
Coupling, standard procedures for SPPS, 16, 244–253
Coupling reagents
  benzotriazol-1-yl-oxytripyrrolidino phosphonium hexafluorophosphate (PyBOP), 13
  O-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluorborate (TBTU), 13
  mechanisms of action, 13–15
  N-[1H-(benzotriazol-1-yl)(dimethylamino)methylene]-*N*-methyl-methanaminium hexafluorophosphate *N*-oxide (HBTU), 13
  chemical structures, 15
  practical guide, 249–250
C-terminal modified peptides, 196–200

Custom peptide suppliers, 168
Cyclic peptide
  automated synthesis, 156–157
  libraries, 151–165
  synthetic strategies, 116–118, 153–154
Cyclosporin A, 151
Cysteine, Npys-protected, 84, 85, 139

D

Dendrimer
  carbohydrate-based, 50
  formation using lipids and sugars, 48–50
  immunogenicity, 48
Dichloromethane (DCM), 4
Difficult sequences, 16–17
Difluoro-phosphonomethyl phenylalanine (F_2PMP)
  F_2PMP-containing dipeptides, 95
  incorporation into peptides, 95–100
  synthesis of N-Fmoc-L-F_2PMP-OH, 94–95
Diisopropylethyalamine (DIPEA), 4
Disulfide bridge or bond formation
  air oxidation, 19, 71
  dimethylsulfoxide (DMSO), 71
  Ellman reagent, 71
  iodine, 71
  redox buffer, 19
  $Tl(tfa)_3$/TFA, 71
DMF, *see* N,N-dimethylformamide
Drug delivery
  using lipid modifications, 44, 47–48
  using sugar modifications, 46–48
  with CPP, 77–89

E

Endothelin-converting enzyme (ECE)
  ECE-1, 144
  peptide substrates, 146
Expressed enzymatic ligation (EEL), of proteins, 123–125
Expressed protein ligation (EPL), 105–130

F

9-Fluorenylmethoxycarbonyl (Fmoc)
  protecting group, 7
  deprotection, 17–18
Fluorescence resonance energy transfer (FRET), 116, 144

G

G protein coupled receptor (GPCR), 30
  chimeric ligand, 31–33
  binding assay, 37–38
Glutathione, 241–243
Glycoprotein D of HSV1, 66
Green fluorescent protein (GFP), 123

H

β-hexoseaminidase, secretion from RBL-2H3 cells, 35–36, 38
High performance liquid chromatography (HPLC)
  analytical, 67
  peptide purification, 67–68, 254–256
High-throughput peptide synthesis, 167–194
Homing sequences, for targeting, 33
Hydrofluoric acid (HF), use in cleavage and deprotection, 67, 251
1-hydroxybenzotriazole (HOBt), 83

I

IMPACT, purification system, 109–113
Intein, 107–109
Intein-mediated purification systems, 111–113

K

Kaiser test for amine determination, 4, 20, 82–83, 244

L

Laboratory guide, for SPPS, 241–257
Ligand binding analysis, 37–38
Linkers, on solid supports for SPPS
  peptide acids, 11
  peptide amides, 10
Lipoamino acids (Laas), 46–47, 52
Lipofection, for intracellular delivery of peptides, 133
Liposaccharides, 48

M

Mass spectrometry
  identification of peptides and proteins, 227–240
  matrix-assisted laser desorption/ionization, time of flight (MALDI-TOF), 245

Mastoparan (MP), 30
MP S, 32
Matrix polymers
polyamide, 8
polyethyleneglycol, 8
polystyrene, 8
Methanol (MeOH), 4
7-Methoxycoumarin-acetate (MCA), 146
Mini-intein, 109

N

Native chemical ligation (NCL), 106
N,N-dimethylformamide, 4
Nuclear localization signal, 135

O

Oligotuftsin, 68–69

P

Parallel synthesis, 168–169
Peptide
arrays, 169
bioconjugates, 209–223
Peptide modification, with lipoamino acids and sugars, 45–61
Peptide nucleic acid (PNA)
antisense applications, 138
cellular delivery, 81, 131–141, 138
conjugation to CPP, 80–81, 134, 137–138
structure, 132
synthesis, 133–134, 135–137
Phosphonomethyl phenylalanine (PMP)
Phosphotyrosine (pTyr)
binding to SH2 domain, 92
binding to PTB domain, 92
structures of mimetics, 92
Phosphotyrosyl mimetics, in signal transduction-directed peptides, 91–103
Piperidine, 4
Protein
cellular delivery with CPP, 77–89
cytotoxic, 121
electrophoresis, 232
engineering, 106–107
enzymatic digestion, for mass spectrometry, 234
epitope tag, 230
fluorescent tag, 106
isolation, 230–231
semisynthesis, 105–130
site-specific modification, 115
splicing elements, *see* intein
staining, mass spectrometry compatible
coomasie, 233
silver, 232
Protein–protein interactions, 122
Proteome, and proteomics, 227–240

Q

Quantitative and qualitative amine tests
bromophenol blue, 257
chloranil, 256
Kaiser, 4, 20, 82–83, 244
picric acid , 256
2,4,6-trinitrobenzene sulfonic acid (TNBS), 21
Quenched fluorescent substrate (QFS)-based peptidase assays, 143–150

R

Reaction vessel, for SPPS, 4, 9
Resins, for SPPS
cleavage with TFA, 18–19
for peptide amide synthesis, 11
for peptide acid synthesis, 11
hydroxymethyl, 11
trityl, 12
RGD sequence, 64

S

Safety-Catch linker, 154
Secretion, of β-hexoseaminidase, 35, 38
Segmental isotopic labeling, for biophysical studies of proteins, 118–121
Selenocysteine, 113
Side-chain protecting groups
common, 5–7, 13–14, 157–158
deprotection, 18–19, 250–251
Npys, 37, 83, 139, 219
Side reactions, during SPPS
aspartamide formation, 20
diketopiperazine formation, 20, 202
prevention with norleucine substitution, 72
Solid phase peptide synthesis (SPPS)
automation, 167–194
basic strategies 3–24
cycles, 6
equipment, 4

high throughput, 167–194
laboratory guide, 241-257
Solid supports, for SPPS, 8–13
Solvents, *see also* individual entries, 9
SPOT synthesis, 170
Split intein-mediated circular ligation of peptides and proteins (SICLOPPS), 118
SPPS, *see* solid phase peptide synthesis
Synthetic antigen, 64

T

Tea-bag synthesis, 170
Tert. Butoxycarbonyl (*t*-Boc)
protecting group, 66–67
synthetic strategy, 66
Thioether bond, 214
Tissue lysis, for protein isolation, 229
Trifluoroacetic acid (TFA), 4,18
Triisopropylsilane (TIS), 4
Tuftsin, 68

V

Vasopressin
antagonist, 32
sequence of [Arg^8]vasopressin, 30
V_{1a} receptor binding assay, 37–38